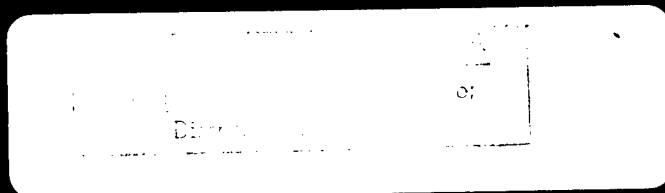# Tight-Binding Approach to Computational Materials Science

EDITORS

Patrice E.A. Turchi

Antonios Gonis

Luciano Colombo

**Tight-Binding Approach
to Computational
Materials Science**

19980721 091

Tight-Binding Approach to Computational Materials Science

Symposium held December 1–3, 1997, Boston, Massachusetts, U.S.A.

EDITORS:

Patrice E.A. Turchi
Lawrence Livermore National Laboratory
Livermore, California, U.S.A.

Antonios Gonis
Lawrence Livermore National Laboratory
Livermore, California, U.S.A.

Luciano Colombo
University of Milan
Milan, Italy

|M|R|S|

CONTENTS

*Invited Paper

PART II: SEMI-EMPIRICAL TIGHT BINDING

*Invited Paper

*Invited Paper

PREFACE

This volume contains the proceedings of the symposium "Tight-Binding Approach to Computational Materials Science", held in Boston, Massachusetts, December 1–3, as part of the 1997 MRS Fall Meeting. Symposium R provided three days of leading-edge research on formal developments in electronic structure studies of materials properties based on the tight-binding approach. The organization of the symposium was motivated both by the usefulness of the tight-binding approximation in the study of the electronic structure of solids, and its ever increasing popularity among materials scientists for simulating properties of complex systems.

The tight-binding model is the simplest scheme within a quantum mechanical framework for describing the energetics of materials which are characterized by fairly localized electrons, such as transition metals and their alloys, or by covalent bonding, such as semiconductors and insulators. Modern tight-binding theory provides a conceptual framework for a physical understanding of the structure of materials and relates the full-scale microscopic, quantum-mechanical computation of materials properties with intuitive chemical and physical arguments. This link between *ab initio* methods and phenomenological concepts allows one to address a wide range of complex materials issues, and at the same time retain the underlying physics responsible for typical materials behavior. Significant efforts were reported at the symposium that improve the computational techniques relying on the tight-binding model in an attempt to bridge efficiently the length and time scales in predicting materials properties in a physically transparent way.

The symposium brought together researchers working on various aspects of tight-binding theory and on its applications to materials science. On the formal front, important inroads were reported in our understanding of first-principles tight-binding methods, the use of tight-binding theory to study the effects of correlations in solids, the development of O(N) methods for electronic structure calculations and molecular dynamics, and parametrization schemes for use with semi-empirical tight-binding models. It was pointed out that electronic structure theory is on the verge of being able to address macroscopic phenomena such as the mechanical properties of metals, using energies obtained from the quantum mechanics of electronic motion. In order to achieve this goal, the calculation of electronic energies for systems with large numbers of inequivalent atoms must be made more efficient than they are at present. Some of the steps for achieving this efficiency based on the use of localized basis sets and the development of new electron energy functionals were reviewed. Recent developments of quantum Monte Carlo in combination with tight-binding models were discussed. Over the last decade, these types of calculations have made the transition from addressing abstract issues concerning the effects of electron-electron correlations on magnetic and metal-insulator transitions, to concrete contact with experiments. Several speakers

discussed ways of obtaining parameters for an implementation of a semi-empirical tight-binding approach to materials science. It was shown that the increased popularity of the tight-binding method is due to the development of linear-scaling methods that allow the use of this formalism in molecular dynamics simulations of large-scale systems.

The application of tight-binding theory to the calculation of materials properties was an important component of the symposium. In this regard, methodology was presented for treating the interaction of light with matter with exciting possible applications to lasers and other systems. As shown by a number of speakers, we know now that tight-binding theory can be generalized to incorporate time-dependent electromagnetic fields in a systematic and gauge-invariant manner. The physics of semiconductors and surfaces, transport in magnetic multilayers, and the properties of thermoelectrics were given ample coverage. Recent developments to calculate conductivities and other linear-response functions, and to provide a new approach to mean-field theories of alloy stability in periodic and nonperiodic systems were reported. This makes possible the prediction of a wider range of materials properties than is hitherto possible. These include optical constants, luminescence in heterostructures, properties in ultra-high magnetic fields, and lattice dynamics in polar materials. Recent progress in the theory of the atomic and electronic structure and magnetism of disordered metallic alloys, surfaces and multilayers was also reported. The predictive power of semi-empirical tight-binding in providing a physical picture of the links between bonding, stability and other properties was illustrated for various systems, such as silicides, carbides, amorphous silicon, nanoclusters and others. The limits of applicability of tight-binding models in the study of surfaces of bimetallic systems were indicated in connection with equilibrium and dynamical processes. A number of applications involving the combination of tight-binding-based electronic structure and molecular dynamics simulations were reported. Finally, the structural properties of elemental covalent liquids, such as arsenic and antimony, as well as amorphous covalent materials, such as Ga-As and Si, can be described accurately through the use of tight-binding simulations. Overall, the symposium discussed in lively and constructive fashion current research issues and the implementation of tight-binding concepts in efficient computational tools which are relevant for addressing the increasingly complex materials challenges of the future, with ample illustrations.

Patrice E.A. Turchi
Antonios Gonis
Luciano Colombo

January 1998

xii

ACKNOWLEDGMENTS

MATERIALS RESEARCH SOCIETY SYMPOSIUM PROCEEDINGS

MATERIALS RESEARCH SOCIETY SYMPOSIUM PROCEEDINGS

Prior Materials Research Society Symposium Proceedings available by contacting Materials Research Society

Part I

First–Principles Tight Binding

THIRD-GENERATION TB-LMTO

O.K. ANDERSEN, C. ARCANGELI, R.W. TANK,
T. SAHA-DASGUPTA, G. KRIER, O. JEPSEN, and I. DASGUPTA.
Max-Planck Institut für Festkörperforschung, Stuttgart, Germany

ABSTRACT

We describe the screened Korringa-Kohn-Rostoker (KKR) method and the third-generation linear muffin-tin orbital (LMTO) method for solving the single-particle Schrödinger equation for a MT potential. In the screened KKR method, the eigenvectors $c_{RL,i}$ are given as the non-zero solutions, and the energies ε_i as those for which such solutions can be found, of the linear homogeneous equations: $\sum_{RL} K^a_{R'L',RL}(\varepsilon_i) c_{RL,i} = 0$, where $K^a(\varepsilon)$ is the screened KKR matrix. The screening is specified by the boundary condition that, when a screened spherical wave $\psi^a_{RL}(\varepsilon, \mathbf{r}_R)$ is expanded in spherical harmonics $Y_{R'L'}(\hat{\mathbf{r}}_{R'})$ about its neighboring sites $\mathbf{R}'$, then each component either vanishes at a radius, $r_{R'}=a_{R'L'}$, or is a regular solution at that site. When the corresponding "hard" spheres are chosen to be nearly touching, then the KKR matrix is usually short ranged and its energy dependence smooth over a range of order 1 Ry around the centre of the valence band. The KKR matrix, $K(\varepsilon_\nu)$, at a fixed, arbitrary energy turns out to be the negative of the Hamiltonian, and its first energy derivative, $\dot{K}(\varepsilon_\nu)$, to be the overlap matrix in a basis of kinked partial waves, $\Phi_{RL}(\varepsilon_\nu, \mathbf{r}_R)$, each of which is a partial wave inside the MT-sphere, tailed with a screened spherical wave in the interstitial, or taking the other point of view, a screened spherical wave in the interstitial, augmented by a partial wave inside the sphere. When of short range, $K(\varepsilon)$ has the two-centre tight-binding (TB) form and can be generated in real space, simply by inversion of a positive definite matrix for a cluster. The LMTOs, $\chi_{RL}(\varepsilon_\nu)$, are smooth orbitals constructed from $\Phi_{RL}(\varepsilon_\nu, \mathbf{r}_R)$ and $\dot{\Phi}_{RL}(\varepsilon_\nu, \mathbf{r}_R)$, and the Hamiltonian and overlap matrices in the basis of LMTOs are expressed solely in terms of $K(\varepsilon_\nu)$ and its first *three* energy derivatives. The errors of the single-particle energies ε_i obtained from the Hamiltonian and overlap matrices in the $\Phi(\varepsilon_\nu)$- and $\chi(\varepsilon_\nu)$ bases are respectively of second and fourth order in $\varepsilon_i - \varepsilon_\nu$. Third-generation LMTO sets give wave functions which are correct to order $\varepsilon_i - \varepsilon_\nu$, not only inside the MT spheres, but also in the interstitial region. As a consequence, the simple and popular formalism which previously resulted from the atomic-spheres approximation (ASA) now holds in general, that is, it includes downfolding and the combined correction. Downfolding to few-orbital, possibly short-ranged, low-energy, and possibly orthonormal Hamiltonians now works exceedingly well, as is demonstrated for a high-temperature superconductor. First-principles sp^3 and sp^3d^5 TB Hamiltonians for the valence and lowest conduction bands of silicon are derived. Finally, we prove that the new method treats overlap of the potential wells correctly to leading order and we demonstrate how this can be exploited to get rid of the empty spheres in the diamond structure.

INTRODUCTION

There is a need for an *intelligible* and *accurate* first-principles electronic-structure method. Our efforts have been directed towards developing a single-particle basis which, for the application at hand, can be adjusted to a useful compromise between being *short ranged, minimal,* and *accurate.*

Recent developments of Multiple Scattering Theory

Since atoms are nearly round, it seems most natural to start out using *spherical waves,* $j_l(\kappa r) Y_L(\hat{\mathbf{r}})$ and $n_l(\kappa r) Y_L(\hat{\mathbf{r}})$, where $L=lm$, as done when solving Schrödinger's equation with the classical multiple-scattering method due to Korringa, Kohn, and Rostoker (KKR).[1] In this method the scattering by the atom at site $\mathbf{R}$ is specified by the *phase shifts,* $\eta_{Rl}(\kappa)$, of its potential well, and the structure of the solid is specified by a Hermitian matrix with elements $B_{RL,R'L'}(\kappa)$, the *structure constants.* In terms of these, the wave-function coefficients, $c_{RL,i}$, are the solutions of the homogeneous, linear equations, one for each $R'L'$:

$$\sum_{RL} [B_{R'L',RL}(\kappa) + \kappa \cot \eta_{Rl}(\kappa) \delta_{R'L',RL}] c_{RL,i} = 0, \tag{1}$$

and the energies, ε_i, are the values of κ^2 ($\equiv \varepsilon$) for which solutions can be found, *i.e.* the determinant of $B(\kappa) + \kappa \cot \eta(\kappa)$ vanishes. There are merely 4–16 equations per atom because all phase shifts with $l > 1$–3 vanish. The KKR equations provide the *exact* solutions of Schrödinger's equation, but only for a muffin-tin (MT) potential, $V(\mathbf{r}) \equiv \sum_R v_R(|\mathbf{r} - \mathbf{R}|)$, which is a superposition of spherically symmetric, non-overlapping potential wells, $v_R(r)$, of ranges s_R.

There are three problems with the KKR method: First of all, a non-overlapping MT potential is a poor representation of the self-consistent potential in any, except the most close packed solid. Secondly, the structure constants have long range, and thirdly, strong energy dependence. Specifically, the energy dependence of the KKR matrix, $B(\kappa) + \kappa \cot \eta(\kappa)$, is not linear like that of the secular matrix, $H - \varepsilon O$, for an energy eigenvalue problem in an energy-independent, possibly non-orthonormal representation. The main reason for the second and third drawbacks is that the spherical Bessel (j_l) and Neumann (n_l) functions have long range and depend on energy. This leads to interferences, which cause long range and strong energy dependence of the structure matrix. For a crystal with lattice translations $\mathbf{T}$, the Bloch-summed structure matrix, $B_{R'L',RL}(\kappa, \mathbf{k}) \equiv \sum_T \exp(i\mathbf{k} \cdot \mathbf{T}) B_{(R'+T)L',RL}(\kappa)$, must be evaluated by the Ewald procedure, has poles at the free-electron parabola, $\kappa^2 = \sum_G |\mathbf{k} + \mathbf{G}|^2$, and a branch cut at the bottom of the continuum, $\kappa=0$.

It was recently shown,[2] and we shall present a slightly different proof below, that even when the potential wells *overlap,* the KKR equations do hold to first order in the potential overlap. This, as we shall demonstrate, allows the use of MT spheres with up to at least 50 per cent radial overlap [$s_R + s_{R'} \lesssim 1.5 |\mathbf{R} - \mathbf{R}'|$ for R and R' denoting nearest neighbors], and hence treat the potential between the atoms in a more realistic way. With such large overlaps, the zero of the potential moves from the potential threshold between the atoms

4

towards the vacuum level, and this means that the energies for the occupied states are usually negative.

It was furthermore shown,[2] that transformation to *linear combinations* of spherical Bessel and Neumann functions, so-called *screened* spherical waves, characterized by a set of background phase shifts $\alpha_{Rl}(\kappa)$, can remove the long range *and* the strong energy dependence from the structure matrix, provided that the energy is not too high. This screening transformation may be expressed as:

$$|n^\alpha\rangle = |n\rangle \left(1 - \frac{\tan\alpha}{\kappa} B^\alpha\right), \text{ where } \tan\eta^\alpha = \tan\eta - \tan\alpha, \text{ and } [B^\alpha]^{-1} = B^{-1} + \frac{\tan\alpha}{\kappa}. \quad (2)$$

These are, respectively, vector-, scalar-, and matrix equations. The superscript labels the representation. In the first equation, we have used a notation in which $|n\rangle$ is a row vector of functions with components $n_l(\kappa r_R) Y_{lm}(\hat{\mathbf{r}}_R)$ and where $\mathbf{r}_R \equiv \mathbf{r} - \mathbf{R}$. The last equation involves inversion of the matrix $B + \kappa \cot\alpha$. For a set of background phase shifts, which are known to give short range, this inversion may be performed in real space and the screened KKR method is basically a first principles tight-binding (TB) method. At the time,[2] however, the relation between range and background phase shifts was poorly understood. This problem was solved later[3] by expressing the background phase shifts in terms of their hard-sphere radii, a_{RL}, defined by

$$\tan\alpha_{RL}(\kappa) \equiv j_l(\kappa a_{RL})/n_l(\kappa a_{RL}), \quad (3)$$

or equivalently, by letting the background phase shifts be those of repulsive potential wells.[4] Looked upon in this way, the role of the confinement is to push the bottom of the continuum up in energy with respect to the floor of the MT potential, and thereby, to leave below a range of energies in which the confined wave-equation solutions are localized and which, in order to be useful, should include the range of the occupied bands for the real (attractive) potential. With the definition (2), the screened spherical waves, $|n^\alpha\rangle$, are still quite energy dependent, but only due to their normalization. A more suitable normalization followed naturally from the hard-sphere point of view.[3]

Finally, it was pointed out[2] that the screening transformation may also be used to remove unwanted channels from the KKR equations by choosing, for those channels, the background phase shifts equal to the real phase shifts, $\alpha(\kappa) = \eta(\kappa)$. This is a transformation to a *minimal basis*.

With these three recent developments, we have the basic ingredients for the Schrödinger part of an *intelligible and accurate* method, the third-generation LMTO method.

Earlier developments; the Atomic Spheres Approximation

The attempts to develop from the KKR method an *intelligible* first-principles method were initiated 25 years ago[5] and overlapping spheres, the two-centre interpretation, and screening transformations[6] have been used routinely for a long time. The new development,[3] which started six years ago,[2] and which will be further elaborated on in the present paper,

aims at making the method also *accurate* without loss of simplicity and elegance. Whereas in the earlier developments reduced range and energy dependence were achieved through a physically motivated approximation, namely the atomic-spheres approximation (ASA), this is now achieved *exactly,* and that in turn, allows a *controlled* approximation for the potential overlap.

The ASA[5,7] consists of letting the MT spheres overlap to the extent that they become space filling, whereby the interstitial region is effectively eliminated so that one may neglect the energy dependence of the wave functions in this region and, hence, the energy dependence of the structure matrix. The remaining energy dependence now occurs only along the diagonal of the KKR-ASA matrix where it enters through the radial logarithmic derivatives evaluated at the atomic sphere,

$$D\left\{\phi_{Rl}\left(\varepsilon, s_R\right)\right\} \equiv s_R \phi'_{Rl}\left(\varepsilon, s_R\right) / \phi_{Rl}\left(\varepsilon, s_R\right). \tag{4}$$

If, as is usually done, the kinetic energy in the interstitial is taken to be zero ($\kappa^2 = 0$), the suitably renormalized KKR-ASA equations become:

$$\sum_{RL}\left[S^0_{R'L',RL} - P^0_{Rl}\left(\varepsilon\right)\delta_{R'L',RL}\right]c_{RL,i} = 0, \quad \text{where}$$

$$P^0_{Rl}\left(\varepsilon\right) \equiv 2\left(2l + 1\right)\left(\frac{w}{s_R}\right)^{2l+1}\frac{D\left\{\phi_{Rl}\left(\varepsilon, s_R\right)\right\} + l + 1}{D\left\{\phi_{Rl}\left(\varepsilon, s_R\right)\right\} - l} \approx \left[\frac{\Delta_{Rl}}{\varepsilon - C_{Rl}} + \gamma_{Rl}\right]^{-1}$$

are the potential functions for well $v_R\left(r\right)$, and C_{Rl}, Δ_{Rl}, and γ_{Rl} are potential parameters. S^0 is the structure matrix given by:

$$S^0_{\text{on site}} = 0, \ S^0_{ss\sigma} = -2\left(w/d\right), \ S^0_{sp\sigma} = 2\sqrt{3}\left(w/d\right)^2, \, , \ S^0_{dd(\sigma,\pi,\delta)} = 10\left(w/d\right)^5\left(-6, 4, -1\right), \tag{5}$$

when we choose $\mathbf{R}' - \mathbf{R} \equiv \hat{z}d$ and w is an arbitrary length scale, usually chosen to be the average Wigner-Seitz radius. The structure matrix thus consists of effective hopping integrals. For monatomic crystals, this gave rise to the concept of canonical bands.[5] However, the $d^{-l-l'-1}$-decay of the hopping integral between orbitals of angular-momentum characters l and l' is too slow for a tight-binding scheme, except for d- and f-orbitals.

It was therefore a breakthrough when it became understood that similarity transformations could be performed on the KKR-ASA ($\kappa^2 = 0$) equations and could lead to short range.[6] The transformation from the bare to a screened representation, specified by screening constants a_{RL}, is given by:

$$P^a\left(\varepsilon\right)^{-1} = P^0\left(\varepsilon\right)^{-1} - a \quad \text{and} \quad \left[S^a\right]^{-1} = \left[S^0\right]^{-1} - a, \tag{6}$$

which are respectively scalar- and matrix equations. a is a diagonal matrix with elements a_{Rl}. The corresponding transformation for the resolvent, useful for Green-function and CPA calculations,[8] is:

$$\left[P^b\left(z\right) - S^b\right]^{-1} = \left(b - a\right)\frac{P^a\left(z\right)}{P^b\left(z\right)} + \frac{P^a\left(z\right)}{P^b\left(z\right)}\left[P^a\left(z\right) - S^a\right]^{-1}\frac{P^a\left(z\right)}{P^b\left(z\right)}. \tag{7}$$

This involves no matrix multiplications, but merely rescaling of matrix elements. As for TB theory, the screening transformation gave a formalism for the "kinetic" part of the often observed dependence of the hopping integrals and on-site elements on the environment.[9] The screening constants yielding short range were found empirically. The potential-dependent choice $a_{Rl}=\gamma_{Rl}$, on the other hand, makes the energy dependence of the KKR-ASA matrix linear (to second order) so that $C + \sqrt{\Delta}S^{\gamma}\sqrt{\Delta} \equiv h^{\gamma}$ becomes the Hamiltonian in an orthonormal, but not necessarily short-ranged basis. The transformation finally made it possible to remove channels from the KKR-ASA equations by choosing $a_{Rl}(\varepsilon) = P^0_{Rl}(\varepsilon)^{-1}$ for such channels.[10] This removal, or "downfolding", however, reintroduced energy dependence of the structure matrix.

When performing density-functional calculations one needs to solve not only Schrödinger's but also Poisson's equation, and with the ASA method this involves approximating not only the potential but also the charge density by a superposition of slightly overlapping, spherically symmetric contributions. This gives a very simple scheme which fails badly in describing total-energy changes caused by symmetry-lowering distortions,[11] however, e.g. the ASA can be used for calculation of pressure-volume relations,[12] but not for calculation of phonon frequencies. Moreover, since the potential spheres are supposed to be space filling in the ASA, open structures can only be treated if the interstices between the atoms are filled with "empty" spheres and this works well only for structures such as the diamond structure, where the interstices have high symmetry. Even in such a case, for the description to be intelligible all empty-sphere channels must be downfolded, and that introduces a rather strong, non-linear energy dependence of the structure matrix which cannot be treated in the LMTO-ASA approach to be discussed below.[10]

Linear Muffin-Tin Orbitals of the first and second generations

In practice one does not solve the KKR equations, but one uses Green functions in a short-ranged representation and at complex energies,[8] or one solves energy eigenvalue equations, $\sum_{RL}[H_{R'L',RL} - \varepsilon O_{R'L',RL}]c_{RL,i} = 0$, which are equivalent with the KKR equations in a certain energy range around some chosen energy, ε_{ν}. In the linear muffin-tin orbital (LMTO) method,[13,5] such an eigenvalue problem is arrived at by using the Raleigh-Ritz variational principle for the Hamiltonian in a basis of LMTO's constructed from the radial Schrödinger-equation solutions, $\phi_{Rl}(\varepsilon,r)$, for the potential wells and their first energy derivatives, $\dot{\phi}_{Rl}(\varepsilon,r)$, at the chosen energy, $\varepsilon=\varepsilon_{\nu}$. In the interstitial region, the first and second generation LMTOs use the spherical waves at κ^2_{ν}, but not their first energy derivatives, so that the energy dependence in the interstitial is suppressed. The second-generation LMTO formalism[6] is elegant, but only in the ASA and only if no channels have been downfolded. Under these conditions, the Hamiltonian and overlap matrices are expressed solely in terms of the structure matrix and the potential functions: The structure matrix enters the formalism in the form of a first-order, two-centre TB-Hamiltonian: $h \equiv \dot{P}^{-1/2}(S - P)\dot{P}^{-1/2}$, where as usual $P(\varepsilon)$ is a diagonal matrix. Here and in the follow-

ing, the common superscript a is dropped and an omitted energy argument means that the energy is set to ε_ν. In terms of this two-centre Hamiltonian the LMTO set may be expressed as $|\chi\rangle \equiv |\phi\rangle + |\dot{\phi}\rangle h$, a form which may be regarded as the matrix equivalent of the linear approximation $\phi_{Rl}(\varepsilon,r) \approx \phi_{Rl}(r) + \dot{\phi}_{Rl}(r)(\varepsilon - \varepsilon_\nu)$. In the basis of these LMTOs the Hamiltonian and overlap matrices are respectively:

$$\langle \chi | -\Delta + V - \varepsilon_\nu | \chi \rangle = h(1 + oh) \quad \text{and} \quad \langle \chi | \chi \rangle = (1 + ho)(1 + oh) + hph, \qquad (8)$$

where

$$o \equiv \langle \phi | \dot{\phi} \rangle = \frac{1}{2!} \ddot{P}/\dot{P} \quad \text{and} \quad p + o^2 \equiv \langle \dot{\phi}^2 \rangle = \frac{1}{3!} \dddot{P}/\dot{P} \qquad (9)$$

are diagonal matrices and it has been assumed that $\phi_{Rl}(r_R) Y_{lm}(\hat{\mathbf{r}}_R)$ is normalized to unity in its sphere, i.e. that $\langle \phi^2 \rangle = 1$. The overlap matrix is seen to be nearly factorized and one may therefore transform to the Löwdin-orthonormalized representation, $|\chi^\perp\rangle \equiv |\chi\rangle \langle \chi | \chi \rangle^{-1/2} \sim |\chi\rangle(1 + oh)^{-1} = |\chi^\gamma\rangle$, in which one finds the following expansion for the Hamiltonian:

$$\langle \chi^\perp | -\Delta + V - \varepsilon_\nu | \chi^\perp \rangle = h - hoh + h\left[oho - (ph + hp)/2\right]h + \dots . \qquad (10)$$

When a gives short range this is a power series in a TB Hamiltonian, h^a. Truncation of this series after the first term yields a spectrum which is accurate in an energy window of size $\sim (10o)^{-1} = \frac{1}{5}\dot{P}/\ddot{P}$ around ε_v, but distorted further away. Adding terms, increases the size of this window at the expense of including further hoppings. The form (10) has been useful in recursion calculations[14] for structurally disordered condensed matter.[15]

For a case like the diamond structure, where one only wants LMTOs centered on atoms, downfolding of the empty-sphere LMTOs is achieved by transformation of the structure matrix using: $\mathbf{a}_E = \mathbf{P}_E^0(\varepsilon_\nu)^{-1}$, with E referring to the empty-sphere channels. The energy is here set to ε_ν because in the LMTO-ASA formalism the structure matrix must be energy independent. Now, an atom-centered LMTO has a tail which extends into the empty spheres, and here, it is substituted by the corresponding partial waves. The atom-centered LMTO is therefore: $|\chi_A\rangle \equiv |\phi_A\rangle + |\dot{\phi}_A\rangle h_{AA} + |\phi_E\rangle \dot{h}_{EA}$, with $\dot{h}_{EA} \equiv \left[-\partial \mathbf{P}_E^0(\varepsilon)^{-1}/\partial\varepsilon|_{\varepsilon_\nu}\right]^{-1/2} \mathbf{S}_{EA}(\varepsilon_\nu) \dot{P}_A^{-1/2}$. This is the way in which the energy dependence of $\mathbf{a}_E(\varepsilon)$ enters, but only to linear order. The overlap matrix $\langle \chi_A | \chi_A \rangle$ will now contain the term $\dot{h}_{AE}\dot{h}_{EA}$ involving $A - E - A$ hoppings, in addition to the terms in (8). This is clumsy and ruins the near factorization of the overlap matrix. With downfolding, the power-series expression (10) for the LMTO Hamiltonian in the Löwdin-orthonormalized basis does therefore not apply.

Most LMTO calculations include non-ASA corrections to the Hamiltonian and overlap matrices, such as the combined correction for the neglected integrals over the interstitial region and the neglected partial waves of high l. This brings in the first energy derivative of the structure matrix, $\dot{S}$, in a way which makes the formalism clumsy.[5,9] Our current, second-generation LMTO code[16] is useful and quite accurate for calculating energy bands because it includes downfolding in addition to the combined correction,[17] but the underlying formalism is so complicated that we never tried to publish it. On the other hand, the combined

8

correction is often important, and so is downfolding because it is the only accurate means of avoiding "ghost bands". The reason for the lost elegance beyond the ASA is that, whereas the LMTO basis is complete to first order in $\varepsilon - \varepsilon_\nu$ *inside* the spheres, it is only complete to zeroth order in the *interstitial*. A compact formalism is therefore obtained only when the interstitial region is neglected, and that is what the ASA does, simply by substituting the MT spheres by space-filling spheres and neglecting the overlap errors. The proof that the KKR equations hold to leading order for overlapping potentials[2] does not apply to the LMTO-ASA formalism.

There *are* LMTO methods sufficiently *accurate* to provide *ab initio* structural energies and forces within density-functional theory.[18] For the reason mentioned above, the LMTOs for such methods[19] are defined with respect to non-overlapping potentials, and since there is considerable probability that a valence or conduction electron is in the interstitial region, *outside* atom-centered, non-overlapping spheres, an accurate basis has to include extra degrees of freedom to describe this region, empty-sphere orbitals centered at interstitial sites and/or atom-centered LMTOs with tails of different kinetic energies (multiple kappa-sets). Moreover, these methods do not use small and short-ranged representations. Finally, since a non-overlapping MT potential is a poor approximation to the self-consistent potential, these methods *must* include the matrix elements of the full potential. Hence, the formalisms are set up to provide final, numerical results and by themselves provide little insight.

Third-generation LMTOs

In this paper we shall modify the LMTO set without increasing its size, in such a way that it becomes complete to first order in the interstitial region too. This is a rather natural thing to do, once the screened spherical waves have been defined in terms of hard-sphere radii. For the MT Hamiltonian, including downfolding, we shall regain the simple formulas from the ASA, provided that $|\phi\rangle$, h, o, and p are suitably redefined. The Hamiltonian and overlap matrices are now given *solely* in terms of the screened and renormalized KKR matrix, which we shall name $K(\varepsilon_\nu)$, and its first three energy derivatives, $\dot{K}(\varepsilon_\nu)$, $\ddot{K}(\varepsilon_\nu)$, and $\dddot{K}(\varepsilon_\nu)$; the potential parameters and the structure matrix do not occur individually as in the formalisms of the previous generations. Third-generation LMTOs[3] thus *do* satisfy the definition that they form a basis constructed to reproduce the wave functions, $\Psi_i(\mathbf{r})$, for a MT potential to linear order in the deviation of the single-particle energy, ε_i, from a freely chosen level, ε_ν. That is, the error of the wave functions is of second order in $\varepsilon_i - \varepsilon_\nu$ and the error of the single-particle energy is then of fourth order. When we use potential wells that overlap, the wave functions will be correct to linear order in the potential overlap and the energy error will be of second order. As we shall demonstrate, this will remedy all shortcomings mentioned above for the previous LMTO generations.

We shall only be concerned with solving Schrödinger's equation in the present paper and leave our LMTO-like expansion of the charge density, solution of Poisson's equation, and evaluation of the total energy and forces for future papers.[20]

We start with a concise yet self-contained derivation of the screened KKR method, which will lead to suitably renormalized versions of Eq.s (2). Then we derive an expression for the error caused by using this method for potential wells which overlap, and find that the error is of second order in the overlap. The weak energy dependence and short range of the screened and renormalized KKR matrix, $K(\varepsilon)$, is exploited by using it to generate few-orbital, low-energy, possibly orthonormal and short-ranged Hamiltonians for a generic high-temperature superconductor (HTSC). Thereafter we derive the new LMTO method and demonstrate by application to free electrons that its energy errors are really of fourth order in $\varepsilon_i - \varepsilon_\nu$. The power and flexibility of the new method is demonstrated by deriving for the HTSC and for diamond-structured silicon various LMTO sets. Using non-orthogonal sp^3 sets for Si, we can get an accurate first-principles description of the valence *and* conduction bands if a 12th-nearest-neighbor range is allowed in the Hamiltonian and overlap matrices. With ε_ν chosen in the middle of the valence band, a 6th-n.n. sp^3-set suffices for an accurate description of the valence band and a reasonable description of the conduction band. In order to halve the number of matrix elements, even for a non-orthogonal basis, we use a formalism analogous to (8) where the off-diagonal elements of o and p have been neglected so that h is the only matrix. With this simplification of the Hamiltonian and overlap matrices, retaining the 6th-n.n. sp^3 basis and the low ε_ν, the description of the valence band remains good and merely the conduction band deteriorates. Finally, it is possible to limit the range to 3rd nearest neighbors provided that d-orbitals are included in the basis. In the last section, we demonstrate that not only for the KKR method, but also for the new LMTO method, the overlap error is of second order and that this can be exploited to get completely rid of the empty-sphere wells in the diamond structure.

SCREENED SPHERICAL WAVES

We start by defining sets of solutions of the wave equation, $[\Delta + \varepsilon]\,\psi(\varepsilon, \mathbf{r}) = 0$, so-called screened-spherical-wave sets, $\{\psi_{RL}^a(\varepsilon, \mathbf{r} - \mathbf{R})\}$, which will serve as interstitial (envelope) functions for the basis that we shall use for solving Schrödinger's equation. The members of a screened-spherical-wave (SSW) set are obtained by letting R run over all atomic sites and L over all angular-momenta for which the scattering is strong. The set is labelled by the superscript a. Instead of defining $\psi_{RL}^a(\varepsilon, \mathbf{r}_R)$ as a specific linear combination of spherical Neumann functions like in Eq. (2), we specify it in terms of an inhomogeneous boundary condition which is illustrated in Fig.s 1 and 2 and is given as follows:

Concentric with each MT sphere, R', we imagine a series of possibly coinciding "*hard*" spheres with radii $a_{R'L'}$. Now, $\psi_{RL}^a(\varepsilon, \mathbf{r}_R)$ is that solution of the wave equation whose $Y_{R'L'}(\hat{\mathbf{r}}_{R'})$ projection on the $R'L'$ sphere equals $\delta_{RL,R'L'}$, that is, 1 on its own sphere and 0 on all other spheres. We do not associate SSWs and hard spheres with *weakly-* and *non-*scattering channels. For such a channel, the $Y_{R'L'}(\hat{\mathbf{r}}_{R'})$ projection of the SSWs is defined to be a regular solution of the corresponding radial Schrödinger equation, that is, it matches onto the irregular wave-equation solution $j_{l'}(\kappa r_{R'}) - \tan \eta_{R'l'}(\kappa)\, n_{l'}(\kappa r_{R'})$, times some con-

stant, $c^a_{R'L',RL}(\varepsilon)$. The weakly- and non-scattering channels are thus parts of the SSW and will not enter the screened KKR- and LMTO matrices explicitly. All high-l' channels are non-scatterers [$\tan \eta_{R'l'}(\kappa) = 0$] due to the dominance of the centrifugal barrier. Empty spheres are examples of a weak scatterers. *Strong* scatterers are then, by definition, those channels with which we associate SSWs and hard spheres. Note that all SSWs in the set have the same boundary condition, except for the $\delta_{RL,R'L'}$. The SSW set, $\{\psi^a_{RL}(\varepsilon, \mathbf{r}_R)\} \equiv |RL\rangle \langle \mathbf{r}|$, may thus be considered as an unperturbed Green function in a hybrid representation.

Fig. 2 shows an SSW for the hypothetical case of only strong scattering. Weak- and non-scattering channels would have shown up as little tails extending into the two hard spheres. Such tails may be seen in Fig. 4 where the dashed curve is an SSW for Si. In this figure we have set the radial functions of the strongly scattering channels to zero inside the hard spheres and, defined in this way, $\psi^a_{RL}(\varepsilon, \mathbf{r}_R)$ jumps by the amount $Y_L(\hat{\mathbf{r}}_R)$ at its own hard sphere, $r_R = a_{RL}$, and has kinks at all hard spheres. Had we instead chosen to continue also the strongly scattering channels of the SSW into the hard spheres, the SSW would have been smooth, but diverging at the sites of the strongly scattering atoms, each radial part going as $j_{l'}(\kappa r_{R'}) - \tan \alpha_{R'L'}(\kappa) n_{l'}(\kappa r_{R'})$. In order to get more feeling for SSWs, let us consider some limiting cases:

If we specify $a_{RL} = a \to 0$ for all channels, we obtain the *bare* spherical waves. These are
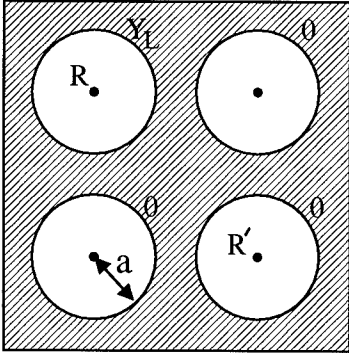
Figure 1: *(above)*. Boundary condition for the screened spherical wave, $\psi^a_{RL}(\varepsilon, \mathbf{r}_R)$. Only strongly-scattering channels are indicated and all hard-sphere radii are equal.

Figure 2: *(right)*. Screened spherical wave centered at the origin, $\psi^a_0(\varepsilon, \mathbf{r})$, and its slopes, $S^a_{R,0}$ *(top)*. The same for its first energy-derivative function, $\dot{\psi}^a_0(\varepsilon, \mathbf{r})$ *(bottom)*.
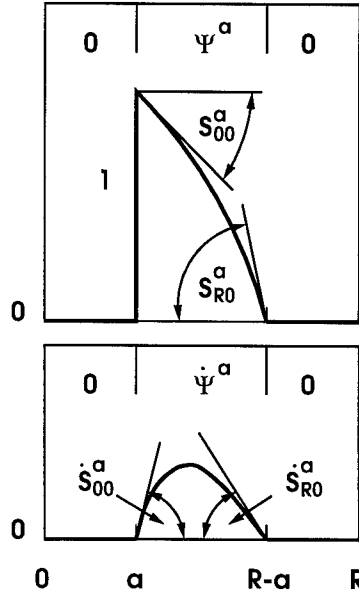
known analytically but, except maybe for small molecules, we never use them. Nevertheless, with the normalization specified above they are:[3]

$$\psi_{0L}^0(\varepsilon, \mathbf{r}) = -\frac{(\kappa a)^{l+1}}{(2l-1)!!} n_l(\kappa r) Y_L(\hat{\mathbf{r}}) = \left[\frac{a}{r}\right]^{l+1} \left[1 + \frac{\varepsilon r^2}{2(2l-1)} - ...\right] Y_L(\hat{\mathbf{r}})$$

$$\to \frac{(\kappa a)^{l+1}}{(2l-1)!!} \frac{\cos(\kappa r - l\pi/2)}{\kappa r} Y_L(\hat{\mathbf{r}}), \quad \text{for} \quad r \to \infty,$$

when $\varepsilon \geq 0$, and where $(2l-1)!! \equiv (2l-1)(2l-3)..\cdot 1$ and $(-1)!! \equiv -1$. When $\varepsilon \leq 0$:

$$\psi_{0L}^0(\varepsilon, \mathbf{r}) = -\frac{(\kappa a)^{l+1}}{(2l-1)!!} [n_l(\kappa r) + ij_l(\kappa r)] Y_L(\hat{\mathbf{r}})$$

$$= \left[\frac{a}{r}\right]^{l+1} \left\{\left[1 + \frac{\varepsilon r^2}{2(2l-1)} -\right] - \frac{(\varepsilon r^2)^l r\sqrt{-\varepsilon}}{(2l+1)[(2l-1)!!]^2} \left[1 - \frac{\varepsilon r^2}{2(2l+3)} +\right]\right\} Y_L(\hat{\mathbf{r}})$$

$$\to \frac{a^{l+1}(\sqrt{-\varepsilon})^l}{(2l-1)!!} \frac{\exp(-r\sqrt{-\varepsilon})}{r} Y_L(\hat{\mathbf{r}}), \quad \text{for} \quad r \to \infty.$$

These expressions hold only for $r \gg a$. Unlike screened spherical waves, the bare ones are eigenfunctions of angular momentum and are independent of the surroundings. For positive energies they have long range. Like all screened spherical waves, the normalization of the bare ones is such that they are dimensionless and, unlike the Bessel and Neumann functions, they depend little on energy near the hard spheres.

Another case is when we specify $a_{RL} = a_R$ for *all* L, and take $\varepsilon = 0$. Then $\psi_{0L}^a(0, \mathbf{r})$ is proportional to the electrostatic potential from a 2^l-pole at the origin, surrounded by grounded conducting spheres with radii a_R centered at the other sites. Since the hard spheres at the neighbors break the spherical symmetry around the origin, the SSW has pure angular-momentum character merely at its own sphere, and this holds only as long as the own sphere coincides with all other spheres concentric with it. Changing the energy will not change the SSW much. If we now let $a_{R'L'}$ be zero for high l''s, the SSWs will "wobble" into the hard spheres.

If the hard-sphere radii are generated from repulsive potential wells,[4] the SSWs are the "impurity states" for that repulsive MT potential.

Since the strongly-scattering components of the SSWs are forced to vanish at all surrounding hard spheres, the *range* of the SSWs depends on the choice of hard spheres and energy: Consider the spectrum ε_i^a of the wave equation with the *homogeneous* boundary condition that the solutions vanish at *all* spheres. This spectrum has a continuum starting at ε_c^a, which in the absence of screening is at zero and which rises with increasing hard-sphere radii. Now, the SSWs are *localized* or *delocalized* depending on whether their energy is below or above the bottom of the continuum. Since we choose energy-independent boundary conditions for the SSWs, their *energy dependence* merely enters through the wave equation, that is through their curvature, and is therefore small when the wavelength exceeds the diameter of the largest interstitial in the hard-sphere solid.

If all hard spheres centered on the same site would coincide, then the hard spheres would have to be *smaller than touching* because, if two spheres had a point (or a circle) in common, then each one of the SSWs centered on the two spheres would be required to be both zero and non-zero at that point (or circle). When only a few low-l channels scatter strongly, neighboring hard spheres may intersect. With decreasing hard-sphere interstitial, the SSW sets thus in general become more and more localized, until the hard spheres start to intersect. Since from there on, the SSWs are forced to change rapidly near the common circles, their behavior becomes chaotic as the circles grow.

We shall generate the screened spherical waves from the bare ones, because those are the only ones we know analytically. Hence, we first consider the question of how to expand an arbitrary wave-equation solution, $\Psi(\varepsilon, \mathbf{r})$, which is regular in all space, except possibly at the atomic sites, in an SSW set, $\{\psi_{RL}^a(\varepsilon, \mathbf{r}_R)\}$, with the same energy. If the number of atoms is finite, this energy is supposed to be negative. Moreover, since $\Psi(\varepsilon, \mathbf{r})$ is a solution of the wave equation, the SSW set is supposed to have no weakly-scattering channels and, *a priory*, we treat all channels as strong scatterers. Finally, we shall not truncate the SSWs inside the hard spheres, but let them continue to the centers. We now expand $\Psi(\varepsilon, \mathbf{r})$ in spherical harmonics on the hard spheres of the SSW set, thus obtaining the coefficients, $\Psi_{Rlm}(\varepsilon, a_{Rlm})$. Unless all of these vanish, the linear combination converges to $\Psi(\varepsilon, \mathbf{r})$:

$$\lim_{\lambda_R \to \infty} \sum_R \sum_{l=0}^{\lambda_R} \sum_{m=-l}^{l} \psi_{Rlm}^a(\varepsilon, \mathbf{r}_R) \, \Psi_{Rlm}(\varepsilon, a_{Rlm}) = \Psi(\varepsilon, \mathbf{r})$$

because, by construction, the linear combination is a solution of the wave equation with the proper energy, and this solution matches $\Psi(\varepsilon, \mathbf{r})$ channel by channel. In order to convince oneself that the latter is sufficient, one may start repeating the argument using an SSW set with L-independent hard spheres. In that case, $\Psi(\varepsilon, \mathbf{r})$ coincides with the linear combination on a closed boundary, because in the case where the system is infinite such a boundary is formed by the entity of all hard spheres, and in the case where the system is finite, the boundary is formed by the hard spheres plus the infinity, where both $\Psi(\varepsilon, \mathbf{r})$ and the linear combination vanish since the energy is negative.

If all the coefficients $\Psi_{Rlm}(\varepsilon, a_{Rlm})$ vanish then $\Psi(\varepsilon, \mathbf{r})$ is an eigenfunction of the hard-sphere solid. In this case a complete set must include, in addition to the SSWs, the degenerate eigenfunctions, or we may choose a different SSW set for the expansion of $\Psi(\varepsilon, \mathbf{r})$.

Changing the hard-sphere radii, but not the sites and the energy, produces another set of SSW's which is also complete in the above-mentioned sense. All such sets are therefore linearly dependent. A set of hard-sphere radii is said to specify a *representation* and the transformation from the a to the b representation is obtained by substituting $\psi_{R'L'}^b(\varepsilon, \mathbf{r}_{R'})$ for $\Psi(\varepsilon, \mathbf{r})$ in the above. Hence, the transformation is

$$\psi_{R'L'}^b(\varepsilon, \mathbf{r}_{R'}) = \sum_{RL} \psi_{RL}^a(\varepsilon, \mathbf{r}_R) \, \psi_{RL,R'L'}^b(\varepsilon, a_{RL}), \tag{11}$$

where $\psi_{RL,R'L'}^b(\varepsilon, a_{RL})$ are the RL components at a_{RL}-spheres of the functions $\psi_{R'L'}^b(\varepsilon, \mathbf{r}_{R'})$.

Expanding now the left and right-hand sides in spherical harmonics on the $b_{R''L''}$-spheres we obtain:

$$\delta_{R''R'}\delta_{L''L'} = \sum_{RL} \psi^a_{R''L'',RL}(\varepsilon, b_{R''L''})\, \psi^b_{RL,R'L'}(\varepsilon, a_{RL}). \tag{12}$$

The two matrices $\psi^a(\varepsilon, b)$ and $\psi^b(\varepsilon, a)$ are thus each others inverses.

SLOPE AND STRUCTURE MATRICES

We have specified the SSWs by their nodes and shall need their radial derivatives at the hard spheres, that is, the dimensionless *slope matrix*. Its element $S^a_{R'L',RL}(\varepsilon)$ is defined as $a_{R'L'}$ times the L'-component of the radial derivative at the $a_{R'L'}$-sphere, with the positive direction taken outwards from $\mathbf{R}'$, of $\psi^a_{RL}(\varepsilon, \mathbf{r}_R)$. This is illustrated in Fig. 2.

In fact knowledge of the hard spheres and the slope matrix makes generation of the SSW set a simple matter: The spherical-harmonics expansion around any site, $\mathbf{R}'$, of any member, $\psi^a_{RL}(\varepsilon, \mathbf{r}_R)$, of the set is given by radial functions and the function for the L' channel is:

$$\psi^a_{R'L',RL}(\varepsilon, r_{R'}) = f_{l'}(\varepsilon, a_{R'L'}, r_{R'})\, \delta_{R'L',RL} + g_{l'}(\varepsilon, a_{R'L'}, r_{R'})\, S^a_{R'L',RL}(\varepsilon). \tag{13}$$

The local expansion converges for $r_{R'}$ smaller than the distance to the nearest site. In (13), f_l and g_l are solutions of the radial wave equation, $[d^2/dr^2 - l(l+1)/r^2 + \varepsilon]\, rf_l = 0$, with the following boundary conditions for $r=a$: f has value one and slope zero, and g has value zero and slope $1/a$.

The SSW-set, $|\psi^a(\varepsilon)\rangle$, may also be expressed *globally* as a linear combination of some known set, $|\psi^b(\varepsilon)\rangle$, as we saw in Eq. (11). The transformation matrix, $\psi^a(\varepsilon, b)$, is then given by Eq. (13) with $r_{R'}$ substituted by $b_{R'L'}$. With the use of Eq. (13), the completeness relation (12) thus expresses the transformation from $S^b(\varepsilon)$ to $S^a(\varepsilon)$.

In order to *generate* the slope matrix, we transform to the bare set, which is known analytically: Using Eq. (13), $S^a(\varepsilon)$ is expressed in terms of $\psi^a(\varepsilon, 0)$, which is computed as the inverse of $\psi^0(\varepsilon, a)$. The latter follows from the local, spherical-harmonics expansion about $\mathbf{R}'$ of the Neuman function centered at $\mathbf{R}$ ($\neq \mathbf{R}'$): $\kappa n_l(\kappa \mathbf{r}_R) Y_L(\hat{\mathbf{r}}_R) = \sum_{L'} j_{l'}(\kappa \mathbf{r}_{R'}) Y_{L'}(\hat{\mathbf{r}}_{R'}) B_{R'L',RL}(\kappa)$, where

$$B_{R'L',RL}(\kappa) \equiv \sum_{l''} 4\pi\, i^{-l+l'-l''} C_{LL'l''}\, \kappa n_{l''}(\kappa |\mathbf{R} - \mathbf{R}'|)\, Y^*_{l'',m'-m}\left(\widehat{\mathbf{R} - \mathbf{R}'}\right) \tag{14}$$

is the KKR structure matrix, which is Hermitian. The summation runs over $l'' = |l' - l|$, $|l' - l| + 2$, ..., $l' + l$, and $i^{-l+l'-l''}$ is real because $C_{LL'L''} \equiv \int Y_L(\hat{r}) Y^*_{L'}(\hat{r}) Y_{L''}(\hat{r}) d\hat{r}$. The on-site elements of $B(\kappa)$ vanish. In this way, we obtain the most important result:

$$aS^a(\varepsilon) - aD\{j(\kappa a)\} = \frac{1}{j(\kappa a)}\left[B(\kappa) + \kappa \cot \alpha(\kappa)\right]^{-1} \frac{1}{j(\kappa a)}, \tag{15}$$

where a, $j(\kappa a)$, $D\{j(\kappa a)\}$, and $\cot \alpha(\kappa)$ are *diagonal* matrices with elements a_{RL}, $j_l(\kappa a_{RL})$, $D\{j_l(\kappa a_{RL})\} = \kappa a_{RL} j'(\kappa a)/j(\kappa a)$, and $\cot \alpha_{RL}(\kappa)$. The quantity $a_{R'L'} S^a_{R'L',RL}(\varepsilon)$, which is $a^2_{R'L'}$ times the L'-component of the radial derivative of $\psi^a_{RL}(\varepsilon)$

at the $a_{R'L'}$-sphere, form the elements of a matrix which is *Hermitian*. This matrix, we call the structure matrix.

For the channels to be treated as strongly scattering with the set $\{\psi^a\}$, we take $\alpha_{RL}(\kappa)$ to be the hard-sphere phase shifts (3), and for those to be treated as weakly scattering and, thus to be *downfolded* into the SSWs, we take $\alpha_{RL}(\kappa)$ to be the real phase shifts, $\eta_{Rl}(\kappa)$. The non-scattering channels do not enter the screening calculation (15), since they neither scatter the bare, nor the screened set. The strongly- and weakly-scattering channels thus contribute to the size of the matrix to be inverted and the strongly-scattering channels are the only ones which will eventually enter the equations for solving Schrödinger's equation.

Instead of expressing (14) and (15) in terms of the usual spherical Bessel and Neumann functions, one could of course have divided the factors κ^l and κ^{-l-1} out on the right-hand side of (15), or used $\psi_l^0(\varepsilon, r)$ instead of $n_l(\kappa r)$, etc.. The only difference between the last equation of (2) and equation (15), is that the Hermitian matrix $aS^\alpha(\varepsilon)$ is normalized in such a way as to make its energy dependence as small as possible, and in such a way as to give $S^\alpha(\varepsilon)$ a geometrical interpretation, namely as the dimensionless slope matrix. Specifically,

$$\kappa^{-1}\tan\alpha\left[B^\alpha(\kappa) - \kappa\cot\alpha(\kappa)\right]\kappa^{-1}\tan\alpha = -j(\kappa a)\,a\left[S^a(\varepsilon) - D\{j(\kappa a)\}\right]j(\kappa a),$$

so that the screened structure matrices $B^\alpha(\kappa)$ and $aS^a(\varepsilon)$ differ because functions of energy have been subtracted from the diagonal elements, and because the rows and columns have been rescaled with such functions. If we form:

$$\mathsf{S}^\mathsf{a}_{R'L',RL}(\varepsilon) \equiv -2\left(w/a_{R'L'}\right)^{l'}\left[S^a_{R'L',RL}(\varepsilon) + (l+1)\,\delta_{R'L',RL}\right]\left(w/a_{RL}\right)^{l+1}, \qquad (16)$$

then $\mathsf{S}^\mathsf{a}(0)$ is the conventional ($\kappa=0$) LMTO structure matrix for the screening constants

$$\mathsf{a}_{RL}(0) = \left[2\left(2l+1\right)\right]^{-1}\left(a_{RL}/w\right)^{2l+1},$$

and $\dot{\mathsf{S}}^\mathsf{a}(0)$ is its first energy derivative for some $\dot{\mathsf{a}}(0)$. For LMTO users who have developed a feeling for the sizes of the conventional structure constants and do not care about the new interpretation in terms of logarithmic derivatives, it is of course possible to use the new method in the conventional "gauge" (16). In that case, one *must* substitute the old potential functions, $\mathsf{P}^\mathsf{a}_{RL}(\varepsilon)$, by $-2\left(w/a_{RL}\right)^{2l+1}\left[D\{\varphi_{Rl}(\varepsilon, a_{RL})\} + l + 1\right]$, with $D\{\varphi_{Rl}(\varepsilon, a_{RL})\}$ evaluated as explained in the following section.[3]

Whereas the slope matrix specifies the normal gradients on the hard spheres of all functions in the SSW set, its first energy derivative, $\dot{S}^a_{RL,R'L'}(\varepsilon)$, specifies the normal gradients of the first-energy derivative functions, $\dot{\psi}^a_{RL}(\varepsilon)$, as illustrated at the bottom of Fig. 2. Since the hard spheres are independent of energy, the energy-derivative functions will vanish at *all* hard spheres, including their own. The first energy derivative of the structure matrix in addition gives the *overlap matrix* of the SSW set: $\langle\psi^a_{RL}(\varepsilon)|\psi^a_{R'L'}(\varepsilon)\rangle = a\dot{S}^a_{RL,R'L'}(\varepsilon)$. This equation follows from the more general one: $\langle\psi^a_{RL}(\varepsilon)|\psi^a_{R'L'}(\varepsilon')\rangle = a_{RL}\left[S^a_{RL,R'L'}(\varepsilon) - S^a_{RL,R'L'}(\varepsilon')\right]/(\varepsilon - \varepsilon')$, which may be derived by use of Green's second

theorem.[3] Here, the strongly scattering radial components have been truncated inside the corresponding hard spheres as illustrated in Fig.s 1 and 2, while the remaining, regular components extend to the centers of the spheres.

Considered as functions of ε, the eigenvalues of the structure matrix $aS^a(\varepsilon)$ have poles when ε coincides with an energy eigenvalue, ε_i^a, of the hard-sphere solid. For practical purposes, the a-radii can be chosen in such a way that the energies of interest to us are well below the bottom of the hard-sphere continuum, and below any localized state of the hard-sphere solid. For such energies, the eigenvalues of $aS^a(\varepsilon)$ are analytical functions of ε. This latter point is demonstrated in the left-hand side of Fig. 3, which also demonstrates that the energy dependence is weak over the ± 10 eV region considered. The right-hand side shows that, for low energies or close sphere packings, the slope matrix decays by an order of magnitude per shell of neighbors. For a monotonically decaying SSW we expect, as illustrated in Fig. 2, a negative slope at its own hard sphere and positive slopes at the neighboring spheres. This is also the behavior found in Fig. 3, at least throughout the first three shells.

In conclusion, the slope matrix generated by inversion of the non singular matrix (15) contains all the information we shall need about the SSW set.
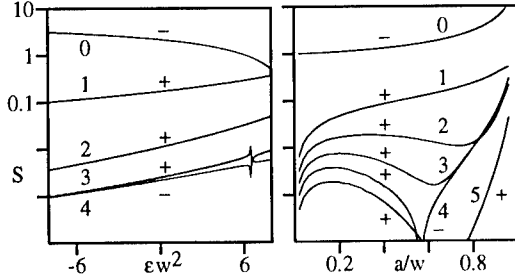
Figure 3: The $ss\sigma$-element, $S^a_{R00,000}(\varepsilon)$, of the slope matrix for the fcc structure and R in the 0'th to 4'th or 5'th shell. *Left:* $a_{spd}=0.7w$. Here w is the Wigner-Seitz radius, which in the fcc structure is 10% larger than that of touching spheres. $\varepsilon w^2=6.05$ corresponds to the lowest free-electron energy at the X point. *Right:* $\varepsilon=0$ and $a_{spd}=a$. The S-scale is logarithmic and the channels with $l>2$ were taken as non-scattering. The number of atoms in the 0'th–5'th shell are respectively 1, 12, 6, 24, 12, and 24. The calculation was performed by matrix inversion, Eq. (15), in real space for a 79-site cluster. For positive energies it was necessary to prevent resonances at the surface of this cluster by enclosing it in a concave sphere simulating the boundary condition $\psi=0$ on the spheres outside the cluster and carrying max(l)=8. The artefact at $\varepsilon w^2 \approx 6.3$ is a surviving resonance. The results are accurate only when the SSW is well localized within the cluster. For cases where the SSW's are not localized within a cluster of affordable size, we must assume crystalline boundary conditions and Bloch-sum $B(\kappa)$ with the Ewald technique before performing the screening inversion (15) for each k-point. With largely overlapping MT spheres, the energies of occupied states are always negative.

SOLVING SCHRÖDINGER'S EQUATION WITH KINKED PARTIAL WAVES

We now come to consider Schrödinger's equation, $[-\Delta + V(\mathbf{r}) - \varepsilon_i]\Psi_i(\mathbf{r}) = 0$, for a MT potential and begin by showing that with our screened spherical waves it is a rather simple matter to formulate the matching problem for the solutions, $\Psi_i(\mathbf{r})$, algebraically:

First we integrate the radial Schrödinger equation for each *strongly* scattering channel *outwards* from the origin to the MT radius s_R in the potential well $v_R(r_R)$, and then *inwards* in *zero* potential (the MT zero) from s_R to the hard-sphere radius a_{RL}. The outwards integration yields the radial partial wave $\phi_{Rl}(\varepsilon, r_R)$, and the subsequent inwards integration yields the radial partial wave "as seen from free space" $\varphi_{Rl}(\varepsilon, r_R)$, with radial logarithmic derivative $D\{\varphi_{Rl}(\varepsilon, a_{RL})\}$ at the hard sphere. These two waves match continuously and differentiably at s_R and they may be seen in the left-hand side of Fig. 4, after multiplication by $Y_L(\hat{\mathbf{r}}_R)$. Let us assume that $\phi_{RL}^a(\varepsilon, r_R)$ and $\varphi_{RL}^a(\varepsilon, r_R)$ have been normalized in such a way that $\varphi_{RL}^a(\varepsilon, a_{RL}) \equiv 1$ at the hard sphere; this is what the superscript a here indicates. In the case where the hard spheres have been chosen to depend on m, radial functions of the same Rl may have different normalizations, hence the subscript L rather than l. With this normalization, the free partial wave matches continuously, but with the kink $S_{RL,RL}^a(\varepsilon) - D\{\varphi_{Rl}(\varepsilon, a_{RL})\}$, to the RL-projection, $\psi_{R'L',RL}^a(\varepsilon, r_{R'})$, of the corresponding SSW, $\psi_{RL}^a(\varepsilon, \mathbf{r}_R)$. Let us furthermore truncate $\phi_{RL}^a(\varepsilon, r_R)$ and $\varphi_{RL}^a(\varepsilon, r_R)$ outside the MT sphere $(0|s_R)$ and, like the SSW, let us truncate $\varphi_{RL}^a(\varepsilon, r_R)$ also inside the a_{RL}-sphere. The function $[\phi_{RL}^a(\varepsilon, r_R) - \varphi_{RL}^a(\varepsilon, r_R)] Y_L(\hat{\mathbf{r}}_R)$ thus equals the proper partial wave inside the hard sphere, where it jumps by $-Y_L(\hat{\mathbf{r}}_R)$, and it vanishes quadratically at the MT sphere with a prefactor proportional to the MT discontinuity $v_R(s_R)$. To this function we now add the corresponding SSW thus obtaining the *kinked partial wave* (KPW):

$$\Phi_{RL}^a(\varepsilon, \mathbf{r}_R) \equiv [\phi_{RL}^a(\varepsilon, r_R) - \varphi_{RL}^a(\varepsilon, r_R)] Y_L(\hat{\mathbf{r}}_R) + \psi_{RL}^a(\varepsilon, \mathbf{r}_R), \tag{17}$$

which is also shown in Fig. 4. This function is everywhere continuous, but has kinks of size $S_{R'L',RL}^a(\varepsilon) - D\{\varphi_{Rl}(\varepsilon, a_{RL})\}\delta_{R'L',RL}$ at the hard $a_{R'L'}$-spheres.

At such a sphere, the kink of the *linear combination* of KPWs, $\sum_{RL}\Phi_{RL}^a(\varepsilon, \mathbf{r}_R)c_{RL}^a(\varepsilon)$, is therefore $\sum_{RL}\left[S_{R'L',RL}^a(\varepsilon) - D\{\varphi_{Rl}(\varepsilon, a_{RL})\}\delta_{R'L',RL}\right]c_{RL}^a(\varepsilon)$. If we can now find an energy, ε_i, and coefficients, $c_{RL,i}^a$, such that

$$\sum_{RL}\left[S_{R'L',RL}^a(\varepsilon_i) - D\{\varphi_{Rl}(\varepsilon_i, a_{RL})\}\delta_{R'L',RL}\right]c_{RL,i}^a = 0 \qquad \text{for all } R'L', \tag{18}$$

then the corresponding linear combination is *smooth* and therefore solves Schrödinger's equation with ε_i as an energy eigenvalue.

The statement that the "kink-cancellation condition" (18) leads to a solution of Schrödinger's equation is exact only for a non-overlapping MT potential. Before continuing to the case of overlapping potentials, let us scrutinize our proof a little closer. Each KPW is constructed to be a solution of Schrödinger's equation at energy ε, except in all shells between concentric MT- and hard spheres, and except for the kinks at the hard spheres. In the case

17

where we choose all concentric hard spheres to coincide with the MT sphere $(a_{RL}=s_R)$, all shells vanish so the statement is obviously true. That it holds also when the hard spheres are different from the concentric MT sphere, follows from the fact that for a linear combination with all kinks cancelled, each $\varphi^a_{R'L'}(\varepsilon_i, r_{R'}) c^a_{R'L',i}$ matches the $R'L'$-projection, $\sum_{RL} \psi^a_{R'L',RL}(\varepsilon_i, r_{R'}) c^a_{RL,i}$, of the linear combination of SSWs, $\sum_{RL} \psi^a_{RL}(\varepsilon_i, \mathbf{r}_R) c^a_{RL,i}$, at $a_{R'L'}$ in value *and in slope*, and since both radial functions are solutions of the *same* second-order differential equation, namely the l''th radial wave equation, they must be *identical*. As a consequence,

$$\varphi^a_{R'L'}(\varepsilon_i, r_{R'}) c^a_{R'L',i} - \sum_{RL} \psi^a_{R'L',RL}(\varepsilon_i, r_{R'}) c^a_{RL,i} = 0, \text{ for } 0 \leq r_{R'} \leq s_{R'} \text{ and } R'L' \in \text{strong scat.}$$

(19)

Inside the $s_{R'}$-sphere then, only terms which satisfy Schrödinger's equation remain, namely $Y_{L'}(\hat{\mathbf{r}}_{R'}) \sum_{RL} \phi^a_{R'L',RL}(\varepsilon_i, r_{R'}) c^a_{RL,i}$ and the weakly- and non-scattering channels of $\sum_{RL} \psi^a_{RL}(\varepsilon_i, \mathbf{r}_R) c^a_{RL,i}$.

The KPW is defined in (17) as the SSW plus the central, pure angular-momentum contribution $\phi - \varphi$, which vanishes quadratically at the MT sphere. In analogy with Slater's augmented plane wave (APW), the KPW might have been named an augmented screened spherical wave. This analogy is only complete though when all hard spheres coincide with their concentric MT sphere.

Next we consider the case of MT *overlap*. Suppose that we have solved the kink-cancellation equations (18) with logarithmic derivatives calculated for potential wells which overlap. To what extent is the resulting *smooth* function, $\Psi_i(\mathbf{r}) \equiv \sum_{RL} \Phi_{RL}(\varepsilon_i, \mathbf{r}_R) c_{RL,i}$, a solution of Schrödinger's equation for the superposition of these overlapping wells? The situation is sketched in Fig. 5 and the answer is, that the smooth superposition of KPWs solves Schrödinger's equation to leading (first) order in the potential overlap.
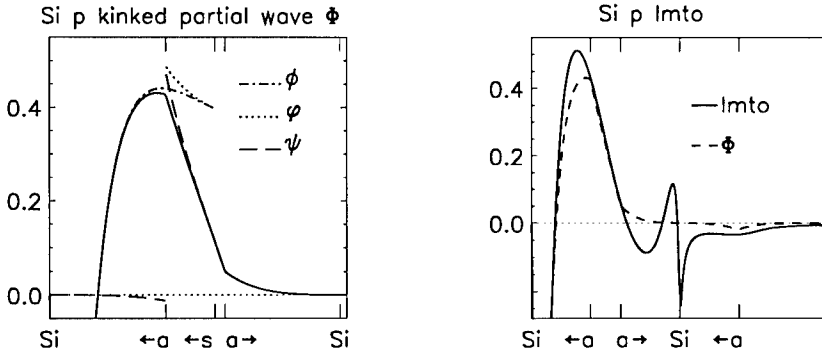
Figure 4: Kinked partial wave (KPW), $|\Phi\rangle \equiv |\phi\rangle - |\varphi\rangle + |\psi\rangle$, and LMTO, $|\chi\rangle \equiv |\Phi\rangle - |\dot{\Phi}\rangle \dot{K}^{-1} K$, for Si p_{x+y+z} plotted along the [111]-direction towards a nearest neighbor in the diamond structure. Note the change of length scale between the left and right panels.

Since we have only considered the strongly-scattering channels in this one-dimensional figure, let us now be a bit more careful. Using the definition (17) and the following definitions: $\psi_i(\mathbf{r}) \equiv \sum_{RL} \psi_{RL}^a(\varepsilon_i, \mathbf{r}_R) c_{RL,i}$, $\phi_R^a(\varepsilon_i, \mathbf{r}_R) \equiv \sum_L \phi_{RL}^a(\varepsilon_i, \mathbf{r}_R) c_{RL,i}$, and similarly for $\varphi_R^a(\varepsilon_i, \mathbf{r}_R)$, we obtain:

$$\left[-\Delta + \sum_R v_R(r_R) - \varepsilon_i \right] \Psi_i(\mathbf{r}) = \sum_{R'} v_{R'}(r_{R'}) \sum_{R \neq R'} \left[\phi_R^a(\varepsilon_i, \mathbf{r}_R) - \varphi_R^a(\varepsilon_i, \mathbf{r}_R) \right]$$

$$+ \sum_R \left[-\Delta + v_R(r_R) - \varepsilon_i \right] \left[\phi_R^a(\varepsilon_i, \mathbf{r}_R) - \varphi_R^a(\varepsilon_i, \mathbf{r}_R) \right] + \left[-\Delta + \sum_R v_R(r_R) - \varepsilon_i \right] \psi_i(\mathbf{r}) =$$

$$\sum_{R'} v_{R'}(r_{R'}) \sum_{R \neq R'} \left[\phi_R^a(\varepsilon_i, \mathbf{r}_R) - \varphi_R^a(\varepsilon_i, \mathbf{r}_R) \right] - \sum_R v_R(r_R) \left[\varphi_R^a(\varepsilon_i, \mathbf{r}_R) - \psi_i(\mathbf{r}) \right] - \left[\Delta + \varepsilon_i \right] \psi_i(\mathbf{r})$$

$$= \sum_{R'} v_{R'}(r_{R'}) \sum_{R \neq R'} \left[\phi_R^a(\varepsilon_i, \mathbf{r}_R) - \varphi_R^a(\varepsilon_i, \mathbf{r}_R) \right] \tag{20}$$

$$= \frac{1}{2} \sum_{R'} v_{R'}(r_{R'}) \sum_{R \neq R'} v_R(s_R) \left[(s_R - r_R)^2 + o\left((s_R - r_R)^2 \right) \right] \phi_R^a(\varepsilon_i, \mathbf{r}_R) \tag{21}$$

$$\sim \frac{1}{2} \sum_{RR'}^{\text{pairs}} v_{R'}(r_{R'}) \left[(s_{R'} - r_{R'})^2 + (s_R - r_R)^2 \right] v_R(s_R) \Psi_i(\mathbf{r}) \tag{22}$$
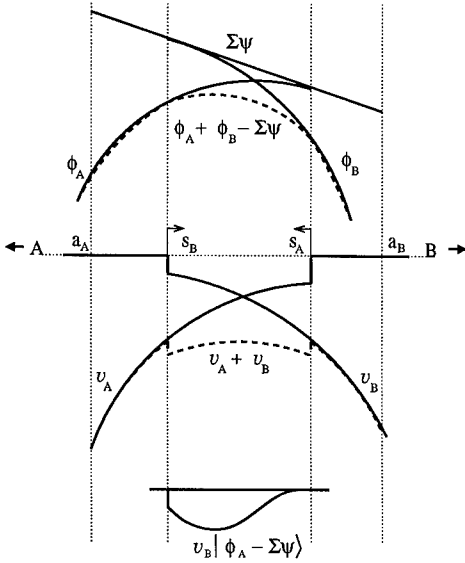
Figure 5: *Middle:* Overlapping potential wells, $v_A(r_A)$ and $v_B(r_B)$, centered at sites **A** (far left) and **B** (far right). *Top:* The solution $\phi_A \equiv \sum_L \phi_{AL}(\varepsilon_i, \mathbf{r}_A) c_{AL,i}$ joins smoothly onto the free solution φ_A at s_A. φ_A runs backwards to a_A, where the kink with the interstitial solution, $\sum \psi \equiv \sum_{RL} \psi_{RL}(\varepsilon_i, \mathbf{r}_R) c_{RL,i}$, is cancelled. Similarly for ϕ_B and φ_B. Due to kink cancellation, the resulting wave function, $\phi_A - \varphi_A + \phi_B - \varphi_B + \sum \psi$, equals $\phi_A + \phi_B - \sum \psi$ in this picture where the angular-momentum character has been suppressed. *Bottom:* The error, $[-\Delta + v_A + v_B - \varepsilon_i] |\phi_A + \phi_B - \sum \psi\rangle = v_B |\phi_A\rangle + v_A |\phi_B\rangle - (v_A + v_A) |\sum \psi\rangle = v_B |\phi_A - \sum \psi\rangle + v_A |\phi_B - \sum \psi\rangle$, consists of two terms each of which, *e.g.* the first, is the product of $v_B(r_B)$ and $\phi_A - \sum \psi$, which vanishes like $v_A(s_A)(s_A - r_A)^2 \phi_A(s_A)$ at the s_A boundary. Hence the error is of *second* order in the potential overlap.

Here, we have first of all made use of the fact that $\Psi_i(\mathbf{r})$ is smooth so that we can apply the Δ-operator to its individual, kinked or discontinuous parts without keeping track of all the resulting diverging terms, because they will cancel in the end. In obtaining the 3rd line, we have used that ϕ solves Schrödinger's equation for its own well. Eq. (20) has then been obtained by use of Eq. (19) for the strongly-scattering partial waves, plus the fact that the weakly- and non-scattering channels (Λ) of $\psi_i(\mathbf{r})$ solve Schrödinger's equation, i.e. that $[\Delta + \varepsilon]\, \psi^a_{RL}(\varepsilon, \mathbf{r}_R) = \sum_{R'} v_{R'}(r_{R'}) \sum_\Lambda \psi^a_{R'\Lambda,RL}(\varepsilon, \mathbf{r}_R)$.

Returning to the strongly-scattering channels of $\sum_R v_R(r_R)\left[\varphi^a_R(\varepsilon_i, \mathbf{r}_R) - \psi_i(\mathbf{r})\right]$, if the overlap is so large that the s_R-sphere overlaps a *neighboring* $a_{R'L'}$-sphere, then it is simplest to imagine that we have *not* truncated the SSWs inside their hard spheres, because otherwise the cancellation (19) would not take place inside the s_R-$a_{R'L'}$ overlap. For consistency then, we should not truncate the free partial waves φ inside their own hard sphere either. The resulting divergencies at the sites, of the SSWs and of the free partial waves, of course cancel for the *smooth* linear combinations. This undoing of the truncation inside the hard spheres is not necessary, but it simplifies the bookkeeping.

The result (20) is then in agreement with what we found in Fig. 5, that the error is a function which vanishes outside the regions of overlap and that inside such a region, it is the product of a function, $v_{R'}(r_{R'})$, which vanishes with a small discontinuity at one of the MT spheres and a function, $\phi_R - \varphi_R$, which vanishes quadratically at the surface of the other MT sphere, with a prefactor proportional to the discontinuity of that MT potential. Remember that the radial part of φ_R is supposed to continue to the origin. The result, which is given in (21), may be obtained from the radial Schrödinger and wave equations. Finally, in expression (22) we have kept only the term of leading order and have used that $\phi^a_R(\varepsilon_i, \mathbf{r}_R) \approx \Psi_i(\mathbf{r})$ in the region picked out by the other factors. Hence, *the error of the wave function is of second order in the potential overlap*.

The error of the one-electron energy may be obtained by first order perturbation theory as: $\Delta\varepsilon_i \equiv \varepsilon_i - \varepsilon_i^{\text{true}} \approx -\langle \Psi_i | -\Delta + \sum_R v_R(r_R) - \varepsilon_i | \Psi_i \rangle$ and, to leading order, we find from Eq. (22) that the error of the band-structure energy is[2]

$$\sum_i^{\text{occ}} \Delta\varepsilon_i \quad \sim \quad -\frac{\pi}{24} \sum_{RR'}^{\text{pairs}} |\mathbf{R} - \mathbf{R}'|^5\, \omega^4_{RR'}\, v_R(s_R)\, v_{R'}(s_{R'})\, \rho\left(\frac{\mathbf{R} + \mathbf{R}'}{2}\right), \tag{23}$$

$$\text{where} \quad \omega_{RR'} \equiv \frac{s_R + s_{R'}}{|\mathbf{R} - \mathbf{R}'|} - 1 \quad \text{is the radial overlap.} \tag{24}$$

In a last section we shall demonstrate how this works for the third-generation LMTO method. Finally, it may be noted that the appearance of the KPW in Fig. 4 is hardly influenced by the MT overlap. This figure in fact applies to an overlap of ω=14%.

Like the slope matrix, the kink matrix is not Hermitian, but the matrix

$$K^a_{R'L',RL}(\varepsilon) \equiv a_{R'L'}\left[S^a_{R'L',RL}(\varepsilon) - D\{\varphi_{Rl}(\varepsilon, a_{RL})\}\,\delta_{R'L',RL}\right] \tag{25}$$

is.[21] This matrix is the renormalized screened KKR matrix. If we multiply each of the kink-cancellation equations (19) with the corresponding hard sphere radius, $a_{R'L'}$, these equations

take the form: $\sum K\mathbf{c} = \mathbf{0}$ and, hence, they are the screened KKR equations. Just as the first energy derivative of the structure matrix is the overlap matrix for the set of SSWs, so the first energy derivative of the KKR matrix is the overlap matrix for the set of KPWs.[22] In fact, one may show that the KKR matrix itself is the energy minus the MT Hamiltonian in the basis of the KPWs with the same energy, that is,

$$K^a_{R'L',RL}(\varepsilon) = \langle \Phi^a_{R'L'}(\varepsilon) \,|\, \varepsilon - (-\Delta + V) \,|\, \Phi^a_{RL}(\varepsilon) \rangle. \tag{26}$$

For Green-function and CPA calculations it has been very important that the transformation (7) of the resolvent, $[P(z) - S]^{-1}$, from one representation to another is merely a scaling rather than a matrix operation. This turns out to hold also in the new formalism, and it means that such calculations may now be performed with more realistic potentials and including downfolding. The result is:

$$K^b(z)^{-1} = a^{-1}g^a(z,b)\,\varphi^a(z,b) + \varphi^a(z,b)\,K^a(z)^{-1}\,\varphi^a(z,b) \tag{27}$$

and has been obtained by use of the completeness relation (12), the one-centre expansion (13), and the following Wronskian relations:[3]

$$ag^b(a) = -bg^a(b), \quad af^b(a) = b^2 g^a(b)', \quad a^2 g^b(a)' = bf^a(b), \quad a^2 f^b(a)' = -b^2 f^a(b)',$$

where the common energy argument, z, has been dropped.

LOW-ENERGY, FEW-ORBITAL, TB HAMILTONIANS; HTSCs

If the energy dependence of the renormalized screened KKR matrix is linearized around some chosen energy ε_ν,

$$K^a(\varepsilon) \approx K^a + (\varepsilon - \varepsilon_\nu)\,\dot{K}^a = -\langle \Phi^a \,|\, -\Delta + V \,|\, \Phi^a \rangle + \varepsilon \langle \Phi^a \,|\, \Phi^a \rangle, \tag{28}$$

then the KKR equations (18) have the form of an algebraic eigenvalue problem. In (28) and in the following, omission of an energy argument ε means that the function is evaluated $at\ \varepsilon_\nu$. The basis set which, by use of the Raleigh-Ritz variational principle for the MT Hamiltonian, gives rise to this problem turns out to be the KPW-set at the fixed energy, ε_ν. This follows from Eq. (26) and is expressed in the second part of Eq. (28). Since the off-diagonal elements of the overlap matrix, $\dot{K}$, only influence the energy eigenvalues to order $(\varepsilon_i - \varepsilon_\nu)^2$, we may even neglect the non-orthogonality of the KPWs and, for a crystal, obtain the correct Fermi surface, $\varepsilon_i(\mathbf{k}) = \varepsilon_F \equiv \varepsilon_\nu$, and the correct group velocities, $\partial\varepsilon_i(\mathbf{k})/\partial\mathbf{k}|_{\varepsilon_F}$, by diagonalization of a first-order Hamiltonian whose matrix elements are simply: $-\tilde{K}^a_{RL,R'L'} \equiv -K^a_{RL,R'L'} \big/ \sqrt{\dot{K}^a_{RL,RL}\dot{K}^a_{R'L',R'L'}}$. This Hamiltonian is completely analogous to h^a in the ASA, but $-\tilde{K}^a$ implicitly contains the integrals over the interstitial region and the downfolded channels, and it works to leading order in the overlap of the potential wells. The range of $-K_{RL,R'L'}$ in R-space, and the size of the energy window inside which the linear approximation holds, depends on the screening. Crudely speaking, the more

strongly-scattering channels included, and the larger their hard spheres chosen without being touching, the shorter is the range of the hopping, and the wider is the energy window.

$-\tilde{K}^a$ can be used as the low-energy, few-orbital, single-particle part of correlated Hubbard-type Hamiltonians, as we shall now demonstrate for a generic high-temperature superconductor (HTSC). We have in the past[23] been able to derive such a Hamiltonian for $YBa_2Cu_3O_7$ using the second-generation LMTO package.[16] That procedure, however, required a lot of hand-work and much insight, and has proved cumbersome to use in general. The new procedure is far more automatic and accurate,[24] and has already proved successful for the ladder compounds.[25]

The basic structural element of all HTSCs is a CuO_2 layer, which is a quadratic lattice with copper at the corners and oxygen halfway between all copper nearest neighbors. In the left-hand side of Fig. 6, the copper sites are those which carry either a $d_{x^2-y^2}$ or an s orbital, and the oxygen sites are those which carry a p_x or a p_y orbital. Different HTSC materials have different stackings of the CuO_2 layers with various "insulating" and/or "doping" layers between them. Nevertheless, the calculated LDA band structures near what is believed to be the Fermi level of optimally doped HTSCs are very similar, and similar to that calculated for the simplest possible such material; dimpled $CaCuO_2$. In this compound, the CuO_2 layers are stacked in the z-direction and are separated by calcium, which sits in the hollow between the eight coppers of the two neighboring layers. The oxygens in the Cu rows running in the x- (y-)direction are dimpled out of the plane by $+(-)$ 7 degrees. The right-hand side of Fig. 6 shows a central CuO_2 layer seen from the side, with a $d_{x^2-y^2}$ and an s orbital on the copper sites and a p_z orbital on the oxygen site. On the CuO_2 layer above is shown a Cu s orbital and on the CuO_2 layer below, an O p_z orbital. Dimpled $CaCuO_2$ is a calculated structure,[26] a theorists dream which hardly exists in this simple form in nature. Its LDA energy bands, which we shall now consider, are nevertheless very similar to those calculated[23] for $YBa_2Cu_3O_7$, one of the only known *stoichiometric* optimally doped HTSCs.

At the Fermi level there is only *one* band per CuO_2 layer, and this is the anti-bonding $pd\sigma$ band formed from the O p_x – Cu $d_{x^2-y^2}$ – O p_y orbitals. This band is at the top of the 10 eV broad O p – Cu d complex consisting of 16 bands, the upper (anti- and non-bonding) part of which may be seen in Fig. 7 (a). According to the LDA and the so-called Van Hove scenario of HTSC, the Fermi level (zero in the figure) for the optimally doped compounds is very close to the saddle-point of the conduction band at $(ak_x, ak_y) = (\pi, 0)$. Hybridization with the Cu s band, which is 5 eV above, has pushed this saddle-point of the anti-bonding $pd\sigma$ band down in energy, to a point where it just "straddles off" the top of the anti-bonding $pd\pi$ bands. This makes the structure susceptible to out-of-row movements of oxygen, because this will mix σ and π bands. In particular the stable structures of $CaCuO_2$ and $YBa_2Cu_3O_7$ have oxygen dimpled seven degrees out of the layer, and this mixes O p_z character into the conduction band in such a way that its saddle-point at $(\pi, 0)$ becomes "extended" that is, the dispersion towards $(0,0)$ becomes proportional to k^4, *i.e. flat*, while in the perpendicular direction, towards (π, π), it remains k^2. The mixing pushes the corresponding $pd\pi$ band

down in energy by about half an eV and leaves the top of other $pd\pi$ bands about an eV below the Fermi level. We thus see that the orbital character of the conduction band, which is the only one we wish to describe, is quite mixed.

The converged LDA bands are showed in panel (a) of Fig. 7. For comparison, panels (b)-(d) show the bands calculated with various "minimal" LMTO sets, specifically, with only the six O p orbitals (b), with only the Cu $d_{x^2-y^2}$ orbital (c), and with the six O p orbitals plus the Cu s and $d_{x^2-y^2}$ orbitals (d). These four calculations all employ the full 3rd-generation LMTO formalism, to be described in the following section, in which the Hamiltonian and overlap matrices, (31) and (30), are given in terms of $K^a(\varepsilon_F)$ and its first three energy derivatives. Panel (b) and (c) demonstrate the power of downfolding in the 3rd-generation LMTO scheme: One may for instance completely leave out the Cu $d_{x^2-y^2}$ LMTOs by attaching that partial-wave character to the tails of the neighboring O p LMTOs (b), or one may completely leave out the O p LMTOs, keeping per cell just the one Cu $d_{x^2-y^2}$ LMTO whose tail then incorporates the O p, Cu s, and other characters (c). As one can imagine, such *massive* downfolding leads to *long* range of the LMTOs. As an example, the Fourier transform of the conduction band shown in panel (c) is the two-centre Hamiltonian in the representation of orthogonalized Cu $d_{x^2-y^2}$ LMTOs, where the cone-like feature of the band around $(0,0)$, caused by near degeneracy of the Cu $d_{x^2-y^2}$ and O p_x orbital energies, gives rise to very long range. This long-ranged, single-band Hamiltonian, we have called (the single-particle part of) the "physical" low-energy Hamiltonian.[23]

What we shall be interested in here is a "chemical" Hamiltonian, which has short range and whose TB parameters behave in a meaningful way when the structure is deformed and
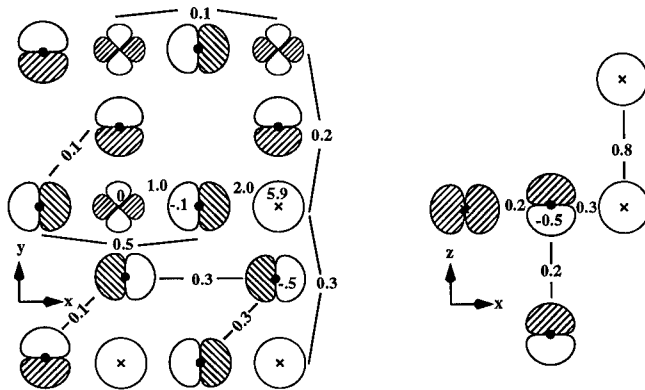
Figure 6: The eight orbitals [O_x $p_{x,y,z}$, O_y $p_{x,y,z}$, Cu s, and Cu $d_{x^2-y^2}$] and the values of their energies (with respect to $\varepsilon_{x^2-y^2}$) and two-centre hopping integrals (eV). These values were obtained as the matrix elements of $-\tilde{K}^a(\varepsilon_F)$ with all channels other than the eight downfolded.

23

when we proceed to similar materials. Which orbitals such a chemical Hamiltonian should contain is then dictated by the range of the corresponding $K^a(\varepsilon)$ matrix. If we imagine a Taylor series like (28), it is conceivable that the higher energy-derivative matrices have longer range. We therefore expect to obtain the shortest range when the energy region of interest is so small that we only need $\tilde{K}^a_{RL,R'L'}(\varepsilon_F)$ as defined above. For dimpled $CaCuO_2$, the chemical basis set turns out to be the one used to generate the bands shown in panel (d). For the same eight orbitals, we show in panel (e) the bands calculated by diagonalization of the effective two-center Hamiltonian $-\tilde{K}(\varepsilon_F)$. We see that this approximation conserves the shape of the conduction band in the relevant range of energy. All computations illustrated
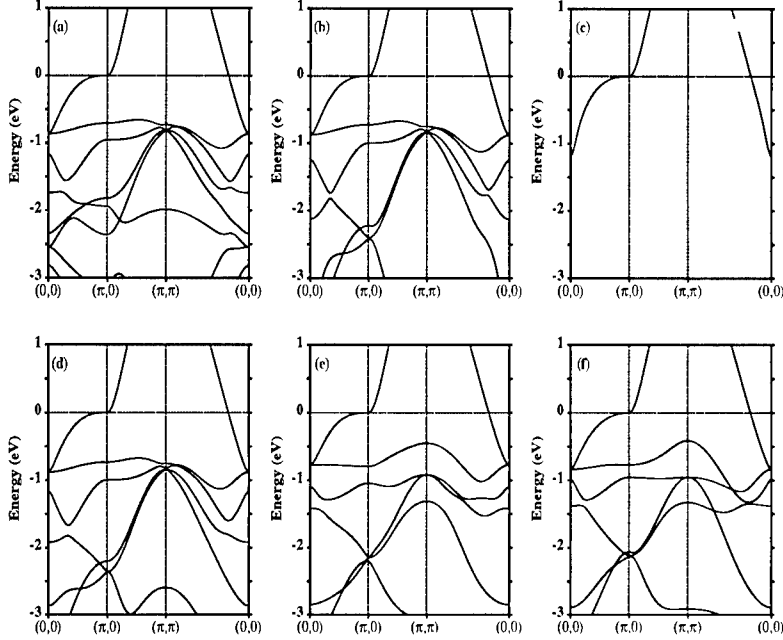
Figure 7: LDA energy bands for dimpled $CaCuO_2$ calculated with the 3rd-generation LMTO method using a converged LMTO basis, *(a)*, and five different simplifications, *(b)-(f)*. Reciprocal-space distances are in units of the reciprocal of the Cu-Cu distance. $\varepsilon_F = 0 = \varepsilon_\nu$. *(a):* All but the $Cu\,s$ and d, and the $O\,p$ orbitals downfolded. $a_{Cu\,s}=a_{Cu\,d}=0.87t$, $a_{O\,p}=0.75t$, where the touching-sphere radius, t, is $1/4$ the Cu-Cu distance. *(b):* All but the six oxygen p orbitals downfolded. $a_{O\,p}=0.96t$. *(c):* All but the single $Cu\,d_{x^2-y^2}$ orbital downfolded. $a_{Cu\,x^2-y^2}=0.62t$. *(d):* All but the six oxygen orbitals, the $Cu\,s$, and the $Cu\,d_{x^2-y^2}$ orbitals downfolded. $a_{Cu\,s}=0.87t$, $a_{Cu\,x^2-y^2}=0.68t$, $a_{O\,p}=0.96t$, *(e):* Like *(d)*, but with $H - \varepsilon_F = -\tilde{K}$ and $O = 1$. *(f):* Like *(e)*, but with $\tilde{K}$ truncated after 3rd-nearest-neighbor hoppings, that is, with the orbital energies and two-centre hopping integrals given in Fig. 6.

24

so far were converged in **R**-space. In fact, they were performed in **k**-space, which means that we started out using the Ewald method to compute $B(\kappa, \mathbf{k})$. When we now Fourier transform $\tilde{K}^a_{RL,R'L'}(\varepsilon_F, \mathbf{k})$, we find that the only non-negligible matrix elements are those given by the orbital energies and two-center hopping integrals in Fig. 6. Panel (f) of Fig. 7 shows the corresponding TB energy bands. This orthogonal, two-center TB Hamiltonian, is seen to reproduce the conduction band very well and to give a satisfactory description of the neighboring bands. This TB Hamiltonian, which we have generated almost automatically, could also have been calculated without the Ewald scheme, by inversion of Eq. (15) in **R**-space. A bit of trial and error is still needed in finding an optimal choice of the hard-sphere radii. The ones we used are listed in the figure caption.

LINEAR MUFFIN-TIN ORBITALS

The first-order Hamiltonian $-\tilde{K}$ does not suffice to describe the energy spectrum over the 10-20 eV range spanned by the valence and lower conduction bands of strongly bonded materials. Nor does inclusion of terms beyond the linear in the Taylor series (28) help, because this does not lead to an algebraic eigenvalue problem. What is needed, is a set of energy independent orbitals which, in contrast to the set of KPWs at a fixed energy, is complete to *linear* order in $\varepsilon - \varepsilon_\nu$.

From a set of KPWs, we first define a set of energy dependent MTOs:

$$|\chi(\varepsilon)\rangle \equiv |\Phi(\varepsilon)\rangle - |\dot{\Phi}\rangle \dot{K}^{-1} K(\varepsilon) \qquad (29)$$

Here and in the following we often drop the common superscript a, and omission of an energy argument means that $\varepsilon = \varepsilon_\nu$. Moreover, we have used the notation in which $|\chi(\varepsilon)\rangle$ is a row vector with elements $|\chi_{RL}(\varepsilon)\rangle \equiv \chi_{RL}(\varepsilon, \mathbf{r}_R)$ and K is a matrix. $\dot{\Phi}_{RL}(\mathbf{r}_R)$ is the first energy derivative at ε_ν of the KPW, $\Phi_{RL}(\varepsilon, \mathbf{r}_R)$, defined in (17). Since the hard spheres are kept independent of energy, the strongly-scattering channels of the energy-derivative functions $\dot{\Phi}$ vanish at *all* the hard spheres. The $\dot{\psi}$-part is sketched in the bottom half of Fig. 2 and the Si p_{x+y+z} MTO at energy ε_ν, that is the LMTO, is shown together with the corresponding KPW in the right-hand side of Fig. 4.

The superposition of $\dot{\Phi}$-functions added to the KPW in (29) is such as to make the MTO *smooth*. That this is so is seen immediately by forming the kink matrix for the MTO: $K(\varepsilon) - \dot{K} \dot{K}^{-1} K(\varepsilon) = 0$. Still, the set of MTOs remains *complete* with respect to the MT potential, because with ε_i being the energy and $\mathbf{c}_i$ a corresponding solution of the KKR equations, $K(\varepsilon_i) \mathbf{c}_i = \mathbf{0}$, we find that the same linear combination of MTOs is: $|\chi(\varepsilon_i)\rangle \mathbf{c}_i = |\Phi(\varepsilon_i)\rangle \mathbf{c}_i = |\Psi_i\rangle$. In contrast to the KPW, the MTO is *independent of energy to linear order* because by differentiation of (29) with respect to energy and subsequent setting $\varepsilon = \varepsilon_\nu$ we get: $|\dot{\chi}\rangle = |\dot{\Phi}\rangle - |\dot{\Phi}\rangle \dot{K}^{-1} \dot{K} = 0$. The energy-independent set of LMTOs, $|\chi\rangle \equiv |\Phi\rangle - |\dot{\Phi}\rangle \dot{K}^{-1} K$, is therefore complete to linear order with respect to the MT Hamiltonian and therefore yields eigenvalues with errors proportional to $(\varepsilon_i - \varepsilon_\nu)^4$. For comparison the conventional single-κ LMTO set is complete to zeroth order in the MT interstitial, albeit to first order in

the spheres, and therefore yields eigenvalue errors of order $(\varepsilon_i - \varepsilon_\nu)^2$ which originate from the interstitial. This is illustrated in Fig. 8. A price for carrying not only ψ, but also $\dot{\psi}$ functions, is that the new LMTO sets corresponding to different hard-sphere radii are no longer linear combinations of each other; the wave-function error, $A^a \cdot (\varepsilon_i - \varepsilon_\nu)^2$, has an a-dependent prefactor.

We now derive the expressions for the Hamiltonian and overlap matrices in the new LMTO basis. For the integrals in all space of KPWs and their first energy derivative functions, one obtains: $\langle \Phi \mid \Phi \rangle = \dot{K}$, $\langle \Phi \mid \dot{\Phi} \rangle = \langle \dot{\Phi} \mid \Phi \rangle = \frac{1}{2!}\ddot{K}$, and $\langle \dot{\Phi} \mid \dot{\Phi} \rangle = \frac{1}{3!}\,\dddot{K}$. The LMTO *overlap matrix* is therefore:

$$
\begin{aligned}
\langle \chi \mid \chi \rangle &= \langle \Phi \mid \Phi \rangle - \langle \Phi \mid \dot{\Phi} \rangle \dot{K}^{-1} K - K\dot{K}^{-1} \langle \dot{\Phi} \mid \Phi \rangle + K\dot{K}^{-1} \langle \dot{\Phi} \mid \dot{\Phi} \rangle \dot{K}^{-1} K \\
&= \dot{K} - \frac{1}{2!}\left(\ddot{K}\dot{K}^{-1}K + K\dot{K}^{-1}\ddot{K} \right) + \frac{1}{3!}K\dot{K}^{-1}\,\dddot{K}\,\dot{K}^{-1}K.
\end{aligned}
\tag{30}
$$

The matrix elements of the MT Hamiltonian used to generate the LMTO set may be found in a similar way. Since the LMTO is smooth there are no problems with Hermiticity like those occurring for the matrix elements between KPWs *alone*. What we mean is, that the result (26) cannot be obtained by naively taking matrix elements of an equation like: $[H - \varepsilon]\,|\Phi\,(\varepsilon)\rangle = 0$, where $H \equiv -\Delta + V$, or of its energy derivative: $[H - \varepsilon_\nu]\,|\dot{\Phi}\rangle = |\Phi\rangle$. For matrix elements between *smooth* linear combinations of KPWs like:

$$
\langle \chi \mid -\Delta + V - \varepsilon_\nu \mid \chi \rangle = \langle \Phi \mid H - \varepsilon_\nu \mid \Phi \rangle - \langle \Phi \mid H - \varepsilon_\nu \mid \dot{\Phi} \rangle \dot{K}^{-1} K - K\dot{K}^{-1} \langle \dot{\Phi} \mid H - \varepsilon_\nu \mid \Phi \rangle
$$
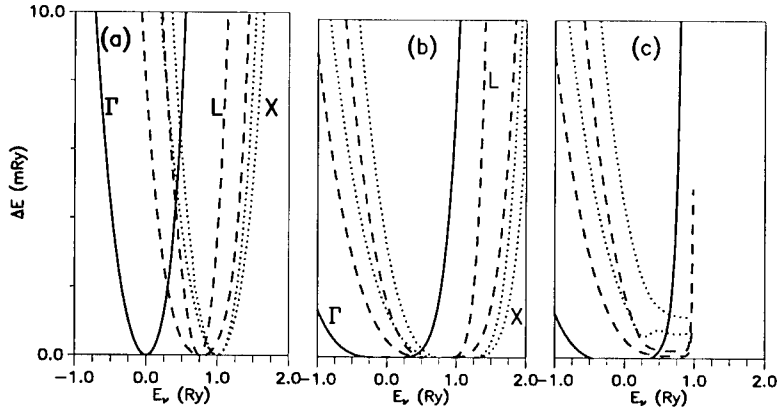
Figure 8: LMTO errors of the lowest free-electron energies (0, 0.75, and 1 Ry) at the fcc Γ, L, and X points as functions of the energy-expansion parameter ε_ν. The states at L and X are doubly degenerate and split when $\varepsilon_\nu \neq \varepsilon$. (a): Old LMTO method with spd-basis and $s=w$. (b): New LMTO method, Eq.s (30) and (31), with spd-basis and $a_{spd}=0.7w=0.77t$; the s-value is irrelevant. In (a) and (b) the bare structure matrix was Bloch-summed with the Ewald technique before it was screened by matrix inversion, Eq. (15), in k-space. (c): Like (b) except that the inversion (15) was performed in R-space using a 79-site cluster enclosed in a concave sphere.

$$+K\dot{K}^{-1}\left\langle\dot{\Phi}\left|H-\varepsilon_\nu\right|\dot{\Phi}\right\rangle\dot{K}^{-1}K$$

$$=-\left\langle\Phi\left|H-\varepsilon_\nu\right|\dot{\Phi}\right\rangle\dot{K}^{-1}K+K\dot{K}^{-1}\left\langle\dot{\Phi}\left|H-\varepsilon_\nu\right|\dot{\Phi}\right\rangle\dot{K}^{-1}K$$

$$=-\left\langle\Phi\mid\Phi\right\rangle\dot{K}^{-1}K+K\dot{K}^{-1}\left\langle\dot{\Phi}\mid\Phi\right\rangle\dot{K}^{-1}K$$

$$=-K+\frac{1}{2!}K\dot{K}^{-1}\ddot{K}\dot{K}^{-1}K \tag{31}$$

such procedures are however correct when used consistently for all terms. Expression (31) thus gives the *MT Hamiltonian matrix* which, together with the overlap matrix (30), are given exclusively in terms of K, $\dot{K}$, $\ddot{K}$, and $\dddot{K}$. These matrices are square and labelled by the channels of the strong scatterers. We stress, that in the 3rd-generation LMTO method, downfolding takes place at the screening stage (15), where it removes the weakly-scattering channels from the structure matrix $S(\varepsilon)$. The calculations for $CaCuO_2$ presented in Fig. 7 (a)-(d) employed this formalism and convincingly demonstrated the new downfolding.

An approximation, which goes beyond the ASA and is not based on dividing space into spheres and neglecting the remainder, consists of neglecting all off-diagonal elements in the *real*-space representation of $\dot{K}$, $\ddot{K}$, and $\dddot{K}$. With this *new* ASA, we have avoided the matrix inversion, $\dot{K}^{-1}$, and the formalism contains only *one* matrix, which we may take to be the first-order two-centre Hamiltonian $-\tilde{K}$ defined in the previous section. This corresponds to renormalizing each KPW and each MTO according to: $\left|\tilde{\Phi}_{RL}(\varepsilon)\right\rangle\equiv\left|\Phi_{RL}(\varepsilon)\right\rangle\Big/\sqrt{\langle\dot{\Phi}_{RL}^2\rangle}=$ $\left|\Phi_{RL}(\varepsilon)\right\rangle\Big/\sqrt{\dot{K}_{RL,RL}}$ and $\left|\tilde{\chi}_{RL}(\varepsilon)\right\rangle\equiv\left|\chi_{RL}(\varepsilon)\right\rangle\Big/\sqrt{\dot{K}_{RL,RL}}$, and the rows and columns of the KKR matrix accordingly: $\tilde{K}_{RL,R'L'}(\varepsilon)\equiv K_{RL,R'L'}(\varepsilon)\Big/\sqrt{\dot{K}_{RL,RL}\dot{K}_{R'L',R'L'}}$. With this renormalization, and taking $\varepsilon=\varepsilon_\nu$, it is easy to see that expressions (30) and (31) reduce to the simple ASA form (8) and (9).

We can develop an exact formalism by Löwdin orthonormalizing the KPWs, instead of merely normalizing them: The overlap matrix for the renormalized KPWs is:

$$\left\langle\tilde{\Phi}_{RL}\mid\tilde{\Phi}_{R'L'}\right\rangle=\dot{\tilde{K}}_{RL,R'L'}\equiv\delta_{RL,R'L'}+\Delta_{RL,R'L'}, \tag{32}$$

where Δ is a Hermitian matrix with vanishing diagonal in RL-space. Its off-site elements $(R\neq R')$ are usually considerably smaller than unity and if we now define a Hermitian matrix: $\dot{\tilde{K}}^{-1/2}=(1+\Delta)^{-\frac{1}{2}}\equiv1-\frac{1}{2}\Delta+\frac{3}{8}\Delta^2-...$, which is the power-series expansion in Δ, then the linear combinations $\left|\bar{\Phi}(\varepsilon)\right\rangle\equiv\left|\tilde{\Phi}(\varepsilon)\right\rangle\dot{\tilde{K}}^{-1/2}$ are seen to form an orthonormal set when $\varepsilon=\varepsilon_\nu$. This is formally like in the conventional ASA. The partial waves truncated outside and normalized inside the atomic s-spheres become in the formalism of the 3rd generation the Löwdin orthonormalized kinked partial waves. The transformed MTO set is: $\left|\bar{\chi}(\varepsilon)\right\rangle\equiv\left|\tilde{\chi}_{RL}(\varepsilon)\right\rangle\dot{\tilde{K}}^{-1/2}=\left|\bar{\Phi}(\varepsilon)\right\rangle+\left|\dot{\bar{\Phi}}\right\rangle h(\varepsilon)$, where

$$h(\varepsilon)\equiv-\dot{\tilde{K}}^{-1/2}\tilde{K}(\varepsilon)\dot{\tilde{K}}^{-1/2}=-\left(1-\frac{1}{2}\Delta+\frac{3}{8}\Delta^2-...\right)\tilde{K}(\varepsilon)\left(1-\frac{1}{2}\Delta+\frac{3}{8}\Delta^2-...\right). \tag{33}$$

Since this expression for the MTO set is also formally identical with an expression which, with the old definitions, was valid only in the ASA, everything else works out the same. *E.g.*, we find: $\left|\dot{\ddot{\chi}}\right\rangle = \left|\dot{\ddot{\Phi}}\right\rangle + \left|\dot{\ddot{\Phi}}\right\rangle h = 0$, because $\dot{h} = -1$. The Hamiltonian and overlap matrices are thus given by (8) with $h \equiv h\left(\varepsilon_\nu\right)$,

$$o \equiv \left\langle \bar{\Phi} \left| \dot{\bar{\Phi}} \right.\right\rangle = \left\langle \dot{\bar{\Phi}} \left| \bar{\Phi} \right.\right\rangle = -\frac{\ddot{h}}{2!}, \text{ and } p + o^2 \equiv \frac{1}{2}\left\langle \bar{\Phi} \left| \ddot{\bar{\Phi}} \right.\right\rangle = \frac{1}{2}\left\langle \ddot{\bar{\Phi}} \left| \bar{\Phi} \right.\right\rangle = \left\langle \dot{\bar{\Phi}} \left| \dot{\bar{\Phi}} \right.\right\rangle = -\frac{\dddot{h}}{3!}. \quad (34)$$

In the 3rd-generation LMTO, h, o, and p are square matrices labelled by the strongly-scattering channels. What we have accomplished is therefore to transform the new Hamiltonian and overlap matrices, (31) and (30), into the form (8), which was previously valid only in the ASA.

In this language the *new* ASA corresponds to neglecting Δ as well as the off-diagonal, real-space parts of o and p. A better approximation is to keep Δ to first order in Eq. (33), and *then* to neglect the off-diagonal parts in the real-space representation of o and p. In this way we still need to specify only *one* matrix, namely the first-order, two-center TB Hamiltonian, h, at the expense of increasing its real-space range somewhat beyond that of $-\tilde{K}$. In the full formalism we have to specify 2 matrices, the Hamiltonian and the overlap matrix or worse, the 3 matrices: h, $\left(\dot{h} = -1\right)$, $\ddot{h}$, and $\dddot{h}$, or even worse, the 4 matrices: K, $\dot{K}$, $\ddot{K}$, and $\dddot{K}$ whose real-space range increases with the number of energy derivatives taken, that is, in order of decreasing importance for the bands near ε_ν.

Some of this is illustrated in Fig. 9 where we compare the LDA band structure obtained from a converged 3rd-generation LMTO calculation (full line) with results (dashed lines) obtained using various *minimal* basis sets, sp^3 in (a)-(c) and sp^3d^5 in (d), and various *truncations*. The empty-sphere *spd*- and, in (a)-(c), the Si *d*-channels were downfolded. Here panel (a) demonstrates that it is possible with merely an sp^3 set to obtain an accurate first-principles description of the valence *and* four lowest conduction bands, provided that we allow the set to be so long ranged that its Hamiltonian and overlap matrices, (31) and (30), extend to 12th-nearest neighbors. This basis is defined by: $a_s=1.1t$, $a_p=1.0t$, and $\varepsilon_\nu = -2$ eV. As usual, t is half the nearest-neighbor distance. If an accurate sp^3 TB-description is needed of merely the valence band, then it is possible to limit the range of the orbitals to the extent that the Hamiltonian and overlap matrices can be truncated after the 6th-nearest neighbors. In (b) this is achieved mainly by shifting ε_ν down to the middle of the valence band. In (c) and (d) we have simplified the calculation of the Hamiltonian and overlap matrices by evaluating (33) to only first order in Δ, and by neglecting the off-diagonal elements in R-space of o and p. As mentioned above, this also makes it necessary to tabulate only *one* two-centre matrix, h. (Note that the *screened* two-centre matrices cannot be completely specified by Slater-Koster two-centre integrals like (5), because the *screened* KPWs and LMTOs do *not* have pure angular-momentum character). Comparison of the dashed lines in (b) and (c) shows that this simplification works for the valence-band structure, but that the quality of the conduction band, which was not aimed at here, has deteriorated.

So far we have not been able with our first-principles procedure to find parameters which will decrease the range of the sp^3 first-order two-centre Hamiltonian, h, below 6th-nearest neighbors. However, with an sp^3d^5 basis this is possible, because then also the d-channels can be used for screening. This is demonstrated in panel (d), where the sp^3d^5-set with the parameters $a_s=a_p=1.0t$, $a_d=0.9t$, and $\varepsilon_\nu = -6$ eV, plus the above-mentioned simplification, yields an h which can be truncated after 3rd-nearest neighbors. The resulting valence band is good and the conduction band very reasonable.

In the past there have been several attempts to model the energy bands of Si by a simple TB Hamiltonian and the need for TB total-energy representations to provide inter-atomic
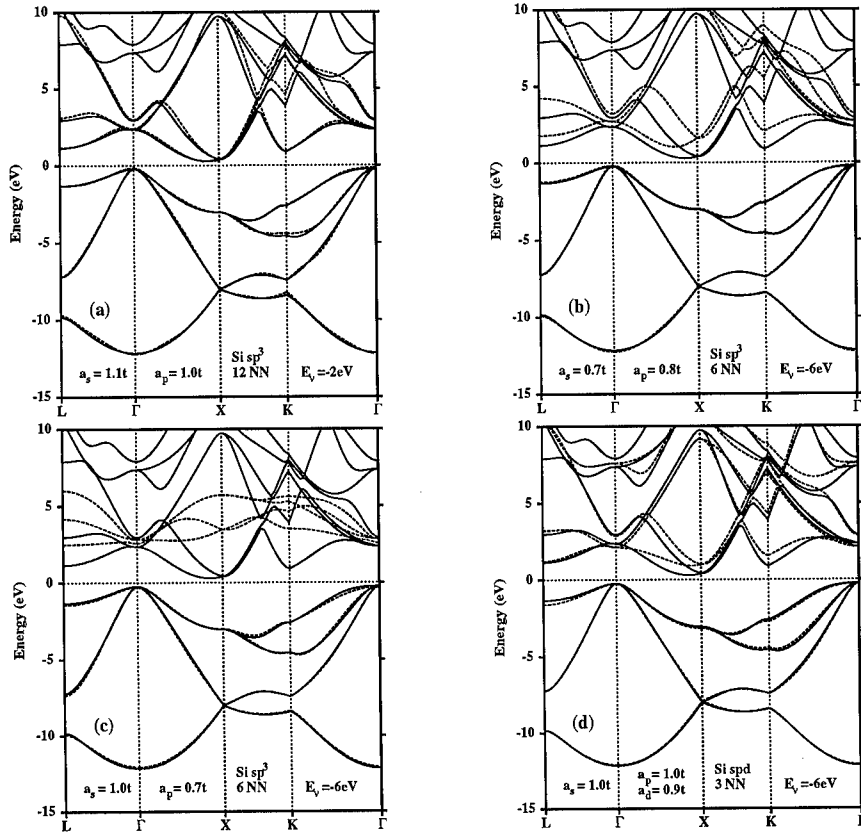
Figure 9: LDA energy bands of diamond structured Si in full lines and various TB approximations thereto in dashed lines (a)-(d). The corresponding LMTO sets and the real-space truncation of the Hamiltonian and overlap matrices (31) and (30) in (a) and (b), and of h (33) in (c) and (d), are specified at the bottom of the panels.

forces for molecular-dynamics simulations has renewed this interest. These attempts range from simple nearest-neighbor, orthogonal parametrization of diamond structured Si with an sp^3 basis in the 70's[27] to recent work with long-ranged non-orthogonal sp^3d^5 basis sets[28] with a hope to provide transferable parameters. All these works relied on fittings of energy bands and total energies obtained from first-principles calculations. Our method is free from such fitting procedures and is purely deterministic. The recent work of McMahan and Klepeis[29] is more similar in spirit to ours, but being based on a full-potential multiple-kappa LMTO calculation with the need for subsequent contraction to a minimal sp^3d^5 basis set, it is more complicated and computationally far more demanding. In fact, our method is so fast, that for us, transferability is no issue. But in all fairness, our total-energy and force calculation is still pending.

GETTING RID OF THE EMPTY SPHERES

The full LDA potential for diamond-structured Si is shown in the top left of Fig. 10. What was used in the LMTO calculations of Fig. 9, however, was the conventional ASA potential shown in the top-right panel of Fig. 10, which is slightly overlapping [ω=14%; see Eq. (24)] and, in addition to the Si-wells, has repulsive wells at the E-sites to describe the hills of the potential. Despite its crude appearance, this ASA SiE-potential, gives nearly exact LDA
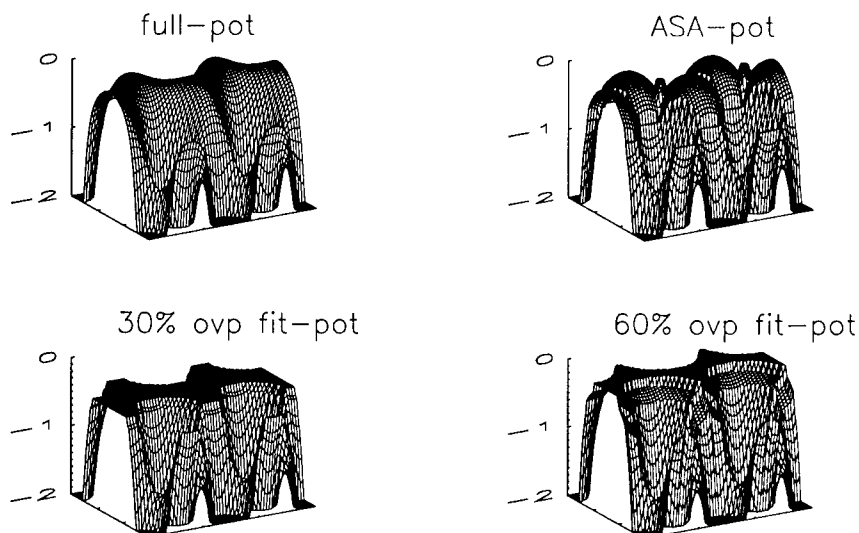
Figure 10: Full LDA potential (in Ry) and various MT approximations for diamond-structured Si in the (110)-plane. The two pseudo potentials at the bottom are least-squares fits to the ASA potential.

valence and conduction bands. But this is a special case. In general, the potential consists of spherically symmetric craters with hills in between, and the latter can be of any shape. Such a potential is naturally modelled by a superposition of atom-centered spherically-symmetric wells, and since we have proved in Eq. (20) that the KKR method can handle such a potential, unless the overlap is too large. The questions are whether this holds also for the new LMTO method, and whether the overlap allowed by these methods is sufficiently large that the MT zero moves up close to the hill tops and the wave functions tail properly off into the voids. Non-MT perturbations would then be local and simple to include. Therefore we first try to treat diamond-structured Si. Two appropriate potentials with respectively 30% and 60% radial overlap are shown in the bottom panels of Fig. 10.

We thus want to fit the full potential, $V(\mathbf{r})$, to a constant (the MT zero) plus a superposition of spherical wells: $V(\mathbf{r}) \sim V_{mtz} + \sum v_R(r_R) \equiv V(\mathbf{r})$. If we decide on a least-squares fit, that is, minimization of $[V - V]^2$, then variation of the functions $v_R(r_R)$ leads to a set of coupled integral equations, one for each R saying that the *spherical* average around site R for radius r_R should be the same for the sum of the MT wells as for the full potential minus V_{mtz}. Variation of V_{mtz} leads to one equation saying that the average of the MT and the full potential should be the same. These equations are fairly simple to solve numerically, but they do not quite express what we want, because a volume element in a region like a void, where the electron has little chance of being, enters with the same weight in the fitting as a volume element in say the bond region. What we really want is a *pseudo* potential which, for a certain band, say the valence band, minimizes the mean squared deviation of the one-electron energies, $\mathrm{Tr}\rho[H - H]^2 = \mathrm{Tr}\rho[V - V]^2$, and this then brings in the electron density, $\rho(\mathbf{r})$, as weighting function. This weighting presents little problem for the δV_{mtz}-equation, which is merely: $\int [V - V]\rho d^3r = 0$, but it complicates the $\delta v_R(r_R)$-equations so much, that we decided on keeping ρ in the δV_{mtz}-equation only. Our MT pseudo potential[30] thus pseudizes the hills rather than the core regions.

Since at this stage, we merely want to see whether we can get rid of the empty spheres
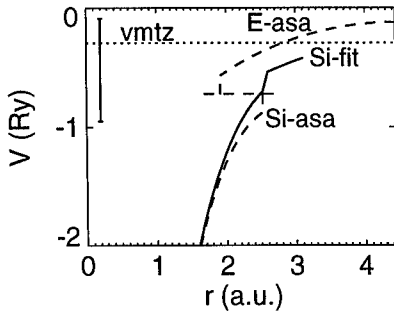
Figure 11: Radial behavior of the Si and E potential wells of the ASA potential (dashed), and of the Si well of the pseudo potential with 40% radial overlap (full). The corresponding 3-dimensional potentials are shown in the upper right and lower left of Fig. 10. The dotted line is the MT zero of the pseudo potential. The vertical bar at the left-hand side indicates the position and extent of the valence band.

in the diamond structure by comparing the valence-band structure calculated for the true potential with that calculated for its pseudo potential, we take the true potential to be one for which we can solve Schrödinger's equation with high accuracy, namely the ASA potential shown in the upper right of Fig. 10. Since this potential is discontinuous at the surfaces of the Si and E spheres, its pseudo potentials, shown at the bottom, are not only discontinuous at s, but also at the Si AS radius. The radial behaviors of the Si and E wells of the ASA potential, as well as that of the Si pseudo potential with 40% radial overlap, are shown in Fig. 11. By comparison of the pseudo potentials with 30% and 60% radial overlap shown at the bottom of Fig. 10, it is obvious that the latter resembles the true potential most closely. Whereas the MT zero of the 14% overlapping ASA potential is only slightly above the bottom of the valence band, that of the 40% overlapping pseudo potential lies 6 eV higher, and that of the 60% overlapping potential is at the top of the valence band.

We have now used the new LMTO method [Eq.s (30) and (31) with a Si sp^3d^5 LMTO set and the Si f-channels downfolded] to calculate the energy bands for the valence-band pseudo potentials as a function of the radial overlap ω. The rms and mean errors of the calculated valence bands are shown in Fig. 12 by diamonds. Since for increasing ω, the potential has increasing range and, hence, increasing freedom, the rms error initially falls, but it eventually rises again as the kinetic-energy errors given by Eq. (20) and proportional to ω^4 take over. The minimum rms error of 80 meV per electron is reached at 30% overlap. The $mean$ error we had expected to vanish for overlaps so small that the kinetic-energy errors are negligible, because the pseudo potential was constructed such that $\int [V - V] \rho d^3r = 0$. Nevertheless, the computation yields a "background" mean error of –50 meV per electron. This is most likely due to errors of second order in $V - V$ caused by the unphysical discontinuities at the E-spheres of the ASA potential. We expect this background error to vanish and thereby the
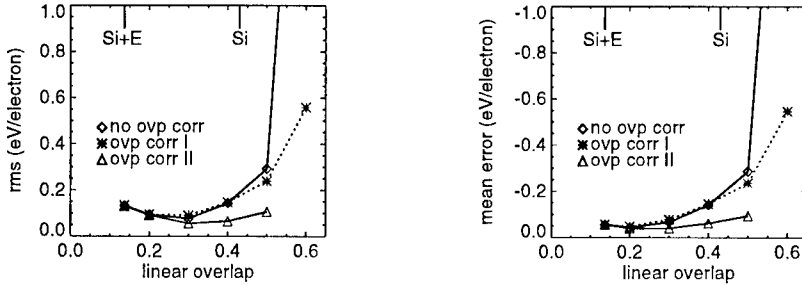
Figure 12: Rms and mean errors of the valence-band energies arising by pseudizing of the ASA potential of diamond structured Si in order to get rid of the E wells. The abscissa is the radial overlap, ω, as defined in Eq. (24). For *space-filling* Si+E spheres, $\omega=14\%$, and for Si spheres 43%. Overlap correction I modifies the pseudo potential, while II includes the proper kinetic energy in the LMTO Hamiltonian.

rms error to be reduced, when for V we use the full potential in the top left panel of Fig. 10.

Since the kinetic-energy error is negative, it represents an attraction between overlapping atoms, and this might cause problems in molecular-dynamics calculations. However, although this attraction increases rapidly with overlap, it does decrease for decreasing interatomic distance and fixed s-radii [see Eq. (23)].

If radial overlaps in excess of $\sim$30 % are needed, then the kinetic-energy error must be corrected. We have tried two schemes, the results of which are given in Fig. 12 by the stars and the triangles. In the first scheme (stars) we have merely modified the pseudo potentials by including in the δV_{mtz}-equation the kinetic-energy error to leading order as given by Eq. (23), whereby this equation becomes: $\int [V - V]\rho d^3 r = \sum_i \Delta\varepsilon_i$. This leads to a reduction of the overlap error, mainly through reduction of the discontinuity $v(s)$. This correction is very simple, but as seen from the figure, hardly sufficient because it only treats the error proportional to $v(s)^2 \omega^4$. Our work on the second scheme (triangles) is still in progress.[30] Here, we evaluate the LMTO Hamiltonian matrix properly to all orders in the overlap, that is, we calculate the LMTO matrix elements following Eq. (20). Of course, this adds terms to expression (31) for the Hamiltonian and spoils the beauty of Eq.s (8), (33), and (34), but we wish to prove that we can control the overlap errors of the new LMTO method, and we want to investigate how large overlaps we can handle. The preliminary results shown in Fig. 12 are encouraging.

REFERENCES

[1] J. Korringa, Physica **13**, 392 (1947); W. Kohn and J. Rostoker, Phys. Rev. **94**, 1111 (1954).

[2] O.K. Andersen, A.V. Postnikov, and S. Yu. Savrasov, in *Applications of Multiple Scattering Theory to Materials Science,* eds. W.H. Butler, P.H. Dederichs, A. Gonis, and R.L. Weaver, MRS Symposia Proceedings No. 253 (Materials Research Society, Pittsburgh, 1992) p 37.

[3] O.K. Andersen, O. Jepsen, and G. Krier in *Lectures on Methods of Electronic Structure Calculations,* edited by V. Kumar, O.K. Andersen, and A. Mookerjee (World Scientific Publishing Co., Singapore, 1994), pp. 63-124.

[4] R. Zeller, P.H. Dederichs, B. Ujfalussy, L. Szunyogh, and P. Weinberger, Phys. Rev. B **52**, 8807 (1995).

[5] O.K. Andersen, Solid State Commun. **13**, 133 (1973); O.K. Andersen, Phys. Rev. B **12**, 3060 (1975); O. Jepsen, O.K. Andersen, and A.R. Mackintosh, Phys. Rev. **12**, 3084 (1975).

[6] O.K. Andersen and O. Jepsen, Phys. Rev. Lett. **53**, 2571 (1984).

[7] H.L. Skriver, *The LMTO Method* (Springer-Verlag, Berlin, 1984).

[8] H.L. Skriver and N.M. Rosengaard, Phys. Rev. B **43**, 9538 (1991); I. Turek, V. Drchal, J. Kudrnovsky, M. Sob, and P. Weinberger, *Electronic Structure of Disordered Alloys, Surfaces, and Interfaces* (Kluwer Academic Publishers, Boston/London/Dordrecht, 1997).

[9] O.K. Andersen, O. Jepsen and M. Sob, in *Lecture Notes in Physics: Electronic Band Structure and Its Applications*, eds. M. Yussouff (Springer-Verlag, Berlin, 1987).

[10] W.R.L. Lambrecht and O.K. Andersen, Phys. Rev. B **34**, 2439 (1986).

[11] O.K. Andersen, Z. Pawlowska and O. Jepsen, Phys. Rev. B **34**, 5253 (1986).

[12] D.G. Pettifor, J. Phys. F **7**, 613 (1977); *ibid.* **7**, 1009 (1977); *ibid.* **8**, 219 (1978).

[13] O.K. Andersen and R.G. Woolley, Mol. Phys. **26**, 905 (1973).

[14] R. Haydock in *Solid State Physics* **35** edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Springer Verlag, Berlin, 1980) p. 129.

[15] H.J. Nowak, O.K. Andersen, T. Fujiwara, O. Jepsen and P. Vargas, Phys. Rev. B **44**, 3577 (1991); P. Vargas C. in *Lectures on Methods of Electronic Structure Calculations*, edited by V. Kumar, O.K. Andersen, and A. Mookerjee (World Scientific Publishing Co., Singapore, 1994), pp. 147-191; S. Frota-Pessoa, Phys. Rev. B **36**, 904 (1987); S.K. Bose, O. Jepsen, and O.K. Andersen, Phys. Rev. B **48**, 4265 (1993).

[16] The Stuttgart TB-LMTO program. http://www.mpi-stuttgart.mpg.de

[17] O. Jepsen and O.K. Andersen, Z. Phys. B **97**, 35 (1995).

[18] O. Gunnarsson, J. Harris, and R.O. Jones, Phys. Rev. B **15**, 3027 (1977); K.H. Weyrich, Solid State Commun. **54**, 975 (1985); M. Springborg and O.K. Andersen, J. Chem. Phys. **87**, 7125 (1986); M. Methfessel, Phys. Rev. **38**, 1537 (1988); M. Methfessel, C.O. Rodriguez, and O.K. Andersen, Phys. Rev. B **40**, 2009 (1989); J. Wills (unpublished); S.Y. Savrasov, Phys. Rev. B **54**, 16470 (1996).

[19] An exception is: L. Vitos, J. Kollar, and H.L. Skriver, Phys. Rev. B **49**, 16694 (1994).

[20] R.W. Tank, O. K. Andersen, G. Krier, C. Arcangeli, and O. Jepsen (unpublished).

[21] In order to make $\dot{K}$ positive-, rather than negative definite, we have defined K with the opposite sign as in Ref. 3.

[22] This neglects the high-order term $\langle \phi - \varphi \mid \psi - \varphi \rangle + h.c. \sim \frac{1}{24} s \left(s - a \right)^4 v \left(s \right) \phi \left(s \right) K$.

[23] O. K. Andersen, O. Jepsen, A. I. Liechtenstein, and I. I. Mazin; Phys. Rev. B **49**, 4145 (1994); O. K. Andersen, A. I. Liechtenstein, O. Jepsen, and F. Paulsen; J. Phys. Chem. Solids **56**, 1573 (1995).

[24] T. Saha-Dasgupta, O. K. Andersen, G. Krier, C. Arcangeli, R.W. Tank, O. Jepsen, and I. Dasgupta (unpublished).

[25] T. F. A. Müller, V. Anisimov, T. M. Rice, I. Dasgupta, and T. Saha-Dasgupta (unpublished) cond-mat/9802029.

[26] S. Y. Savrasov and O. K. Andersen, Phys. Rev. Lett. **77**, 4430 (1996).

[27] D. J. Chadi and M. L. Cohen, phys. status solidi **68**, 405 (1975).

[28] R. E. Cohen, L. Stixrude, and E. Wasserman, Phys. Rev. B **56**, 8575 (1997).

[29] A. K. McMahan and J. E. Klepeis, Phys. Rev. B **56**, 12 250 (1997).

[30] C. Arcangeli, O.K. Andersen, and R.W. Tank (unpublished).

EFFICIENT ELECTRONIC ENERGY FUNCTIONALS FOR TIGHT-BINDING

Roger Haydock
Physics Department and Materials Science Institute, University of Oregon, Eugene, OR 97403-1274

ABSTRACT

Generalized functionals are constructed from the exchange-correlation energy by a Legendre transformation which makes the new functionals stationary at the electronic charge density, potential, and wave functions for the ground-state. Using generalized functionals, the density, potential, and wave functions can be independently parameterized and varied to determine the ground-state energy-surface for a system of atoms. This eliminates the computationally awkward steps of constructing densities from wave functions or potentials from densities, and is particularly well suited to parameterizations using tight-binding orbitals together with atomic-like densities and potentials. For each choice of parameters, the only quantities which must be computed are the electron-electron energy for the density, the integral of the potential over the density, and the band structure energy for the wave functions. To second order in the density, potential, and wave functions, the energy for a configuration of atoms is given by the generalized functional evaluated at a superposition of atomic densities, a potential made by stitching together the atomic potentials where they are equal, and atomic wave functions. For more accurate stationary energies the densities, potentials, and wave functions can be improved by one or more conjugate gradient steps.

BIG SYSTEMS, TIGHT-BINDING, AND THE HOHENBERG-KOHN FUNCTIONAL

Electronic structure theory is on the verge of being able to address macroscopic phenomena such as the mechanical properties of metals, using energies obtained from the quantum mechanics of electronic motion. In order to achieve this goal, the calculation of electronic energies for systems with large numbers of inequivalent atoms must be made even more efficient. Some of the steps toward achieving this efficiency are the use of localized basis sets and the development of new electronic energy functionals.

The purpose of this paper is to explain how electronic energy functionals other than the Hohenberg-Kohn (H-K) functional can be constructed and applied to localized bases, here loosely called tight-binding. This Sec. outlines the difficulties of using plane wave basis sets and of the self-consistent minimization of the H-K functional. There follow Secs. describing the construction of generalized functionals with convenient variational parameters, the calculation of energy differences and forces from the generalized functionals, the construction of approximate electron densities and potentials together with their optimization, and a brief discussion of the advantages of atomic exchange and correlation over the Local Density Approximation (LDA).

The reason it is necessary to use localized electronic bases for big systems is to keep small the number of non-zero matrix elements in each row and column of the electronic Hamiltonian. Atomic-like orbitals for the electrons only extend over a few shells of neighbors, so the calculation of Hamiltonian matrix elements between localized orbitals involves only a few atoms, and there are at most a few hundred non-zero elements in each row or column of the Hamiltonian matrix. With localized bases the effects of distant parts of the system arise from the way the localized orbitals combine to form extended states. Unlike localized orbitals, each plane wave overlaps every atom in the system, so calculation of the Hamiltonian requires the Fourier transform of the potential, a process whose length increases faster than the number of inequivalent atoms. Worse, the number of non-zero matrix elements in each row or column of the Hamiltonian grows as the number of inequivalent atoms making any attempt to diagonalize the Hamiltonian grow as the cube of the number of non-zero elements.

The use of localized bases is further justified by a remarkable property of wave equations such as the electronic Schroedinger equation, that the density of states, weighted by its intensity on some localized function, is exponentially insensitive to distant parts of the system [1]. For the

present discussion, this property tells us that the energy it takes to displace an atom or the force on an atom, is insensitive to distant parts of the system. This means that plane wave bases with their equal weighting of all parts of the system are ill suited to the calculation of quantities such as energy differences and forces due to local disturbances, which depend only on nearby parts of the system.

The energy differences and forces are just the quantities which should be calculated, not total energies as is often done. The simplest way to see this is that as the number of inequivalent atoms grows, so does the total energy, while the energy differences and forces do not. Consequently, the total energies have to be calculated ever more accurately in order to keep constant the accuracy of the energy differences. Furthermore, even the energy difference between two structures can be calculated as a sum of localized changes in the structure. Insensitivity of the local electronic structure to distant parts of the system also justifies abandoning the total energy approach.

For the last thirty years, electronic energies have largely been calculated by self consistent solution of the Kohn-Sham (K-S) equations, which arise from minimizing the H-K electronic energy functional. The H-K theorem [2] states that the energy of the electronic ground state is a unique functional of the electron density $n(r)$. The usefulness of this theorem comes from the K-S minimization [3] of the electronic energy functional $E[n(r)]$ which is,

$$E[n(r)] = T[n(r)] + <V_i(r) \, n(r)> + U_{ee}[n(r)], \tag{1}$$

where $T[n(r)]$ is the minimal kinetic energy for electron density $n(r)$, $<V_i(r) \, n(r)>$ is the ionic potential energy of $n(r)$, and $U_{ee}[n(r)]$ is the minimal electron-electron interaction energy for the electron density, including both the Hartree energy and the exchange and correlation energy which must be approximated in practice. Note that as used here $U_{ee}[n(r)]$ is not local because of its Hartree contribution, although the exchange and correlation energy in $U_{ee}[n(r)]$ is local in LDA.

The H-K functional cannot be minimized directly by variation of the electron density because the kinetic energy functional $T[n(r)]$ cannot be evaluated directly; it is itself the minimum expectation value of the kinetic energy operator T over all the wave functions which have the given density. Instead, this functional must be minimized by setting its gradient to zero, subject to the constraint that the integral of the electron density remain constant. This produces the K-S equations for a set of functions $\{\Psi_\alpha(r)\}$ and energies $\{E_\alpha\}$,

$$(T + V_i(r) + V_{ee}(r)) \, \Psi_\alpha(r) = E_\alpha \Psi_\alpha(r), \tag{2}$$

where $V_{ee}(r)$ is the functional derivative of the electron-electron interaction energy with respect to the electron density, which is just the sum of the Hartree and exchange-correlation potentials of the electrons. The ground state electron density $n(r)$ is then the sum over electrons of the squares of the $\Psi_\alpha(r)$ which are each occupied by a pair of electrons in order of increasing E_α, and the kinetic energy is the sum over electrons of the expectation values of T for the occupied functions. Approached this way, minimizing the H-K functional is equivalent to solving Eq. 2 which is non-linear because $V_{ee}(r)$ depends on the functions which satisfy the equation. Equation 2 is so complicated, it is really only practical to solve it by iteration to self consistency.

The first of the three problems with the self consistent solution of Eq. 2 is that within one self consistent cycle, a charge density must be constructed from the $\{\Psi_\alpha(r)\}$ which satisfy Eq. 2. Since this charge density is a combination of products of functions, it has many more degrees of freedom than the functions themselves, forcing further approximations. The second problem is that in order to construct the electron-electron potential $V_{ee}(r)$ from $n(r)$, the functional derivative of $U_{ee}[n(r)]$ must be evaluated This requires at least the solution of Poisson's equation for the Hartree part of $V_{ee}(r)$ and can requires a third basis in which to expand $V_{ee}(r)$, different from the other two bases and usually truncated. On top of the problems of different bases, solutions of Eq. 2 are sensitive to $V_{ee}(r)$ and so the truncation errors in the two additional bases compound in the solution of Eq. 2.

What is needed is an electronic energy functional which can be evaluated directly without intermediate minimizations of complexity comparable to the original problem, and without the

need to solve enormously complicated non-linear equations. This is just what the use of generalized functionals accomplishes. The idea for generalizing the H-K functional is to decouple the natural variables from one another to produce functionals which can be optimized without the construction of charge densities from wave functions or the solution of Poisson's equation, and without resort to self consistency.

GENERALIZED FUNCTIONALS

This derivation of generalizations of the H-K functional begins with the more intuitive expression for the electronic energy as the band structure energy from which the electronic interaction energy must be subtracted because it is included twice in the band structure energy. The band structure energy is the sum over electrons of the energies $\{E_\alpha\}$, obtained by solving Eq. 2, of the occupied functions, and the interaction energy includes the Hartree energy of the electrons as well as their exchange and correlation energies. While it is natural to think of the interaction energy as depending on the electron density $U_{ee}[n(r)]$, the band structure energy is more naturally related to the electronic potential $B[V_{ee}(r)]$. However, in the spirit of H-K, take the interaction potential at each point to be a functional of the electron density $V_{ee}(r)[n(r)]$, namely the functional derivative of the interaction energy with respect to the electron density. But for the fact that the band structure energy minus the interaction energy is not stationary at the ground state density, the electron energy would be $B[V_{ee}(r)[n(r)]]-U_{ee}[n(r)]$, where the band structure energy is a functional of the interaction potential which is in turn a functional of the electron density.

The problem with the stationary point of the above functional can be understood by examining the changes in band structure and interaction energies produced by small changes in the electron density. The change in band structure energy δB for a small change in electron density $\delta n(r)$ is $<|\Psi|^2(\delta V_{ee}/\delta n)\,\delta n(r)>$ where $(\delta V_{ee}/\delta n)\,\delta n(r)$ is the change in the interaction potential due to the change in electron density and $|\Psi|^2$ is the sum over electrons of the squared magnitudes of the occupied functions. The change in interaction energy is $<V_{ee}(r)[n(r)]\,\delta n(r)>$, and is independent of the change in the band structure energy because it does not depend on $|\Psi|^2$, so stationary points of this functional do not coincide with those of the H-K functional.

The fault can be corrected by a Legendre transformation which changes variation in $V_{ee}(r)$ to variation in $n(r)$, namely the subtraction of $<V_{ee}(r)n(r)>$ from the band structure energy. This subtraction reduces the band structure energy to the sum of the kinetic and ionic energies, to which the interaction energy must be added to get the correct electronic energy. The generalized functional is now,

$$F[n(r)] = B[V_{ee}(r)[n(r)]] - <(\delta U_{ee}/\delta n)[n(r)]\,n(r)> + U_{ee}[n(r)], \qquad (3)$$

which consists of a band structure energy minus the double-counting term, the latter consisting of $<(\delta U_{ee}/\delta n)[n(r)]n(r)>-U_{ee}[n(r)]$ because of the Legendre transformation. This functional is often called the Harris-Foulkes functional [4, 5] although is has been used by others including [6, 7]. In Eq. 3 the coefficient of δn vanishes when $|\Psi|^2$ is equal to $n(r)$ which is just the self consistency condition for the K-S equations. It is important to note that the cancellation of variations with respect to the density is implicit so that $|\Psi|^2$ need never be constructed.

The generalized functional is stationary rather than extremal at the ground state energy and density, see Ref. [5], so that energies obtained from it are neither upper nor lower bounds on the exact energy. Although the H-K functional is minimal at the ground state, the usual approximations for the exchange and correlation energies are not variational, so the K-S approach also lacks bounding properties.

This generalized functional still has the fault that the interaction potential $V_{ee}(r)$ has to be constructed from the electron density by solving Poisson's equation. As discussed in the context of the H-K functional, this requires a basis which is independent of the basis used to calculated the band structure energy, and the band structure energy is particularly sensitive to errors in the

37

construction. The advantage of this functional over the H-K functional is that the electron density does not have to be constructed from the solutions to Eq. 2.

Pursuing the approach of tailoring the functional to the needs of the computation, the above functional can be further generalized and improved by decoupling the electronic potential from the electron density. As stated above, the band structure energy is naturally a functional of the electronic potential, and expressing it that way generates a generalized functional of two variables, the potential and the density,

$$G[n(r), V_{ee}(r)] = B[V_{ee}(r)] - <V_{ee}(r) n(r)> + U_{ee}[n(r)]. \qquad (4)$$

The stationarity condition for $G[n(r), V_{ee}(r)]$ is that both its linear variation with respect to the electron density and with respect to the interaction potential be zero. In this case the change in the band structure energy is just $<|\Psi|^2 \delta V_{ee}(r)>$ which cancels the variation of $<V_{ee}(r) n(r)>$ with respect to $V_{ee}(r)$ only when $|\Psi|^2$ is equal to $n(r)$, and the variation of $<V_{ee}(r) n(r)>$ and $U_{ee}[n(r)]$ with respect to $n(r)$ only cancel when $V_{ee}(r)$ is the functional derivative of the interaction energy. These two stationarity conditions are just the relations between $|\Psi|^2$ and $n(r)$, and between $V_{ee}(r)$ and $n(r)$ in the K-S equations.

The functional $G[n(r), V_{ee}(r)]$ avoids the difficulties of the K-S equations and of $F[n(r)]$ because $|\Psi|^2$ need never be constructed, and $V_{ee}(r)$ is not constructed from $n(r)$. The potential and the density are independent, so there are no errors introduced by incompatibility in their representations, although there are still the truncation errors due to finite representations of each quantity. Furthermore the independence of the density and potential prevents errors in the density from affecting the band structure energy.

It is possible to add further arguments to the generalized functionals, for example the electronic functions $\{\Psi_\alpha(r)\}$ [8]. Determination of the band structure energy for $G[n(r), V_{ee}(r)]$ still requires the solution of Eq. 2 for the given interaction potential. In terms of a basis, this is a matrix diagonalization which can be approached as a minimization of the band structure energy with respect to the electronic functions to produce a functional whose arguments are $\{\Psi_\alpha(r)\}$, $n(r)$, and $V_{ee}(r)$ to give,

$$\Phi[\{\Psi_\alpha(r)\}, n(r), V_{ee}(r)] = \sum_{occ} <\Psi_\alpha(r)^*(T + V_i(r) + V_{ee}(r)) \Psi_\alpha(r)>$$

$$- <V_{ee}(r) n(r)> + U_{ee}[n(r)]. \qquad (5)$$

Variation of Φ with respect to $\{\Psi_\alpha(r)\}$ vanishes for the $\{\Psi_\alpha(r)\}$ which satisfy Eq. 2, and the other variations vanish as for the functional $G[n(r), V_{ee}(r)]$.

The functional $\Phi[\{\Psi_\alpha(r)\}, n(r), $ and $V_{ee}(r)]$ has the advantage that its evaluation does not require solution of any equations, just the evaluation of matrix elements and integrals. It is also well adapted to optimization procedures because its variation about its stationary point is the most nearly quadratic of all the functional discussed so far. Its variation in $\{\Psi_\alpha(r)\}$ is strictly quadratic and its variation in $V_{ee}(r)$ is strictly linear. The only non-quadratic variation is in $n(r)$ and arises from the dependence of the interaction energy on $n(r)$.

The difficulty with the functional $\Phi[\{\Psi_\alpha(r)\}, n(r), V_{ee}(r)]$ is that the number of degrees of freedom for the functions $\{\Psi_\alpha(r)\}$ scales with the system size(number of inequivalent atoms). There is a function for each pair of electrons, so the number of functions scales linearly with the system size, and if the functions are delocalized, then the number of components of each function also scales linearly with the system size making the total number of components of all the occupied functions quadratic in the system size. Since there is a parameter for each component of the electronic functions, the computational effort required to evaluate the functional increases quadratically with system size when the electronic functions are delocalized. The functional $\Phi[\{\Psi_\alpha(r)\}, n(r), V_{ee}(r)]$ is most useful when the electronic functions can be localized, for example

38

when bands are either empty or full.

The electron density $n(r)$ and the interaction potential $V_{ee}(r)$] each have a number of components which grows linearly with system size because there is only one of each function in the cell containing each atom. The only question about the linear scaling of $G[n(r), V_{ee}(r)]$ is whether the band structure energy can be evaluated with a computation scaling linearly with system size. Density matrix methods attempt to do this, but the recursion method discussed below takes advantage of the fact that changes in local energies are insensitive to distant parts of the system, to make possible calculations of energy differences which are independent of the system size.

PARAMETERIZATION

Although the generalized functionals avoid the problems of solving non-linear equations and of constructing electron densities as well as potentials, there remain a large number of integrals which must still be evaluated. The most obvious of these integrals is the expectation value of the interaction potential for the charge density $<V_{ee}(r) \, n(r)>$ but similar integrals occur in the evaluation of the interaction energy $U_{ee}[n(r)]$, and in the evaluation of matrix elements of the kinetic energy T, ionic potential $V_i(r)$, and interaction potential $V_{ee}(r)$ needed to calculate the band structure energy. The accuracy of the electronic energy increases with the number of degrees of freedom included in variations of the electron density and interaction potential, but so does the number of integrals, so it is well worth expanding the electron density and interaction potential and electronic functions in a way which makes the integrations efficient.

In applications of the generalized functional to the calculation of defect energies in MgO[9], the electronic functions were expanded in Slater orbitals for atoms and in Gaussians for vacancies, the electron density was expanded in the squares of these orbitals, and the potentials were expanded in Laguerre polynomials for the screened Coulomb potential. These choices were convenient although the calculation of three-center integrals, integrals of two orbitals with different centers over a potential with a third center, dominated the computational effort because there were so many of them and because each was a long computation.

The importance of three-center integrals in many systems of interest and the computational effort they require, encourage a search for more efficient representations of the orbitals and potentials. One group of candidates are localized orbitals which tridiagonalize the kinetic energy operator [10]. In this approach the potentials are expanded as operators in outer products of the orbitals, thereby converting the three-center integrals in expectation values to linear combinations of overlap integrals between basis orbitals. These overlaps are two-center integrals whose angular dependence can be expressed using Slater-Koster [11] parameters, and whose distance dependence is easily parameterized.

It may be possible to make large parts or even the whole of electronic energy calculations analytic by inspired choices of parameterization for which the integrals have analytic expressions.

ENERGY DIFFERENCES AND FORCES

As is mentioned above, the greater the system size, the larger the magnitude of the total energy, and the greater the accuracy to which the total energy must be calculated to maintain a fixed accuracy in differences in total energies. If the structure changes only locally such as when a vacancy or interstitial forms, there are cancellations between the total energies before and after the change which leave their difference relatively small. It is crucial to carry out these cancellations implicitly for systems with many inequivalent atoms because the total energies get arbitrarily large compared to the energy differences.

One approach to making implicit the cancellations in the total energies is to calculate forces, which reduce to calculating the expectation value of the derivative of the ionic potential with respect to the change in atomic positions, because first derivatives of other contributions cancel when the functional is stationary. If only a few atoms move, then the derivative of the ionic potential is localized near the atoms, and electronic functions are only needed in a small region of space. The problem with forces is that energy differences are integrals over the forces, so they have to be calculated at many intermediate steps as an atom is moved, for example, to make a vacancy. Each time the force is to be calculated, the functional has to be optimized for the new

structure and the calculation grows with the number of steps.

Another approach[12] to calculating energy differences uses the recursion method to make the cancellations implicit. Instead of forces, the element of these calculations is the change in band structure energy when a particular orbital is removed from all the occupied electronic functions. If $u_0(r)$ is the orbital to be removed, successive orbitals $u_1(r)$, $u_2(r)$, ..., $u_N(r)$,are constructed by recursion [13] to tridiagonalize Eq. 2. The difference between the sum of occupied eigenvalues of this tridiagonal matrix, with and without the row and column corresponding to $u_0(r)$, is the difference in band structure energy.

Because Eq. 2 is a wave equation, its local densities of states have the black body property of being independent of distant parts of the system. As a result, the differences in band structure energy calculated this way converge for small tridiagonal approximations of dimension 10 to 20, independent of the size of the system. In other words, by this method, differences in band structure energies can be computed with an effort which is independent of the number of inequivalent atoms.

An energy difference such as a vacancy energy is made up of the change in the band structure energy, taking into account the changes in the ionic and interaction potentials, and the change in the double counting terms which depend on the change in the electron density and the interaction potentials. The first step is to calculate the initial and final densities and potentials by optimizing the energy functional. As long as the change in the electron density does not affect distant parts of the system significantly, the changes in the electron density and potentials need be considered only in a region containing $u_N(r)$ which contributes to the change in the band structure energy.

The difference in band structure energies between removing one orbital with one energy and removing another orbital with a different energy is the same as replacing the first orbital with the first energy by the second orbital with the second energy. To calculate the band structure energy of the vacancy, add up the energy differences for removing each orbital of the vacancy atom and its matrix elements for the initial potentials, and subtract the energy differences for removing each orbital of the vacancy and its matrix elements for the final potentials. The rest of the vacancy energy is due to the change in the double counting term $<V_{ee}(r)\, n(r)> - U_{ee}[n(r)]$ which can be calculated from just the initial and final densities and potentials.

Even energy difference between two phases can be calculated from local structural changes similar to the formation of a vacancy. The two phases are arranged with an interface between them, and the interface contains a typical site for the growth of one phase from the other, see further Gibson and Haydock [12]. The energy to transform one chemical unit from the first to the second phase is the sum of energy differences as successive atoms at the growth site move from their positions in the first phase to their positions in the second phase. At no time during this does a total energy have to be calculated, and the accuracy of the total energy difference is only slightly reduced from the accuracy of the energy differences for moving each atom. Because the growth site is typical, its change in energy is typical, and the total change in energy as one phase grows into the other along the interface is just the energy difference of the typical unit times the number of units.

The important conclusion of this Sec. is that calculations of differences in total energy can always be formulated as a local change in the structure, and that the computation required scales linearly with the number of atoms displaced, but is independent of system size rather than linear in the system size as one might think. This conclusion holds ever if there are long-ranged changes in the potential due to a local change in the electron density, because sufficiently far from where the electron density changes, the effects of the change in potential can be calculated using the dielectric constants of the material.

While energy differences are adequate for most purposes including the relaxation of structures, forces can also be calculated in the above approach. When the functional is stationary, the only non-zero contribution to the derivative of the total energy with respect to atomic displacement is the derivative of the band structure energy with respect to the ionic potential, so instead of the differences in the energies of the occupied eigenvalues of the tridiagonal matrix, we need the derivatives of the energies of the occupied eigenvalues with respect to the ionic potential. The force comes out to be the sum over electrons of the derivatives of the occupied eigenvalues of the tridiagonal matrix with respect to the ionic potential, which is better computed from these eigenvalues rather than by integrals over densities of states.

DENSITIES, POTENTIALS, AND OPTIMIZATION

The main purpose of the generalized functionals is to get electronic energies directly from structures, with the absolute minimum of computation. The generalized functionals are good for this because they give electronic energies which are accurate to second order in the errors of their arguments, the electron density, interaction potential, or even the wave functions. The stationarity property also means that any effort applied to improving the inputs is repaid by a squared improvement in the energies, so there is a big incentive to construct good inputs.

The construction of an electron density for a solid starts with atomic electron densities because binding energies in solids are small compared to the energy it takes to change the electron density inside an atom. The best approach is to take the atomic electron density which would lie outside the atomic cell and redistribute it in the bonding regions near the surface of the atomic cell, but it is often adequate simply to overlap the atomic electron densities, even though this slightly changes the density deep in the atom where the density should not change.

Interaction potentials constructed from overlapped atomic electron densities as in [4] usually contain serious errors because the overlapping of atomic densities alters the densities deep inside the atoms. This happens because the change in the interaction potential due to a small change in the electron density is large where the electron density is large, namely deep inside atoms. A big change in the potential produces big changes in the electron functions which lead to big changes in both the kinetic and ionic contributions to the band structure energy, contributions which are not canceled by the subtraction of the integral of the electron density over the interaction potential.

It is much better [14] to construct interaction potentials for the solid directly from interaction potentials for the atoms by stitching them together on the surfaces, between atoms, where they are equal - make the interaction potential at each point in the solid equal to the atomic interaction potential which has the greatest magnitude at that point. The advantages of this construction are that atomic interaction potentials are known to great accuracy from fits to atomic levels, the construction is simple to carry out, and errors introduced by overlapping atomic electron densities are much reduced.

The generalized functional $\Phi[\{\Psi_\alpha(r)\}, n(r), V_{ee}(r)]$ is particularly useful if approximate localized electronic functions are available. Just as for the construction of approximate electron densities and interaction potentials, a good guess for the electronic functions is the valence orbitals of the atoms. Even if the electron functions are delocalized, local energies depend only nearby atoms, so approximate electronic functions need only be good in the vicinity of the disturbance whose energy is to be computed.

If construction of the electron density and interaction potential from atoms does not give sufficient accuracy, then the stationary property of the generalized functionals can be used to improve the result. The simplest approach is conjugate gradients [15] which determines the stationary point in a finite number of steps if the variation around the point is quadratic, but works well near any non-singular stationary point. Electronic energy functionals can be singular where levels cross in the band structure energy, and where the exchange and correlation energies are singular in the electron density, but these singularities can be removed by partial occupations of levels and smoothing of the exchange and correlation energies.

A conjugate gradient optimization begins by choosing a promising search direction; for calculation of an energy difference due to a local structural distortion, the search direction should be a component of the electron density or interaction potential which is expected to couple strongly to the distortion. The coefficient of the component is varied to find its stationary value (zero derivative), holding other parameters constant. The next search direction is the linear combination of the current search direction and the gradient of the functional at the stationary point of the current search, such that the derivative of the functional along the new search direction is conjugate (orthogonal) to the current search direction. This conjugacy of search direction means that as the density or potential is varied along the new search direction, the functional remains stationary with respect to variation along the current search direction.

The general step of the conjugate gradient optimization begins with a search direction which is conjugate to all previous search directions. The functional is varied to find the stationary point; its gradient is evaluated at the stationary point; and a new conjugate search direction is

constructed as a linear combination of the gradient, the current search direction, and the previous search direction so that the gradient of the functional along the new search direction is orthogonal to the current and previous search directions. The property of conjugate gradients which makes them efficient is that by making the new search direction conjugate to the current and previous directions, it is also made conjugate to all previous search directions near a quadratic stationary point.

In addition to the functional itself, conjugate gradient optimization uses the gradient of the functional which is the vector whose components are derivatives of the functional with respect to each of its parameters, the parameters for the electron density, interaction potential, or even the electron functions. The derivative of the generalized functional $G[n(r), V_{ee}(r)]$ with respect to the coefficient of one component of the interaction potential is the difference between the expectation value of that component for the occupied electronic functions, and the integral of that component over the electron density. The derivative of the generalized functional with respect to the coefficient of one component of the electron density is the integral of the that component over the difference between the interaction potential obtained from the functional derivative of the interaction energy, and the interaction potential which is the argument of the functional. The gradient of the generalized functional is made up of the energy expectation values of the inconsistencies in the components of electron density and interaction potential. The gradient of

$\Phi[\{ \Psi_\alpha(r)\}, n(r), V_{ee}(r)]$ with respect to the electronic functions can be calculated using the same ideas.

Since calculations of energy differences and forces require accurate electron densities and interaction potentials only near where the structure is changing, only the needed components of these quantities need be optimized. Again, the computation required is independent of the size of the system.

THE WIGNER-SEITZ EXCHANGE AND CORRELATION

LDA approximates exchange and correlation energies in solids with those of the homogeneous electron gas, and this approximation begins to break down when charge densities in the solid vary significantly over the Fermi wavelength of the homogeneous gas. Electron densities near atoms can vary enormously, so it is not surprising that LDA fails to get band gaps accurately. Logic suggests that approximations for exchange and correlation based on atoms should do much better.

An example of an atomic approximation for exchange and correlation is the Wigner-Seitz trick [16, 17, 18] for alkali metals in which the electronic potential at each point in space is that of the nearest atom, *not* a superposition of the potentials of nearby atoms. The idea behind this is that the electronic and ionic potentials of neighboring atoms cancel. This trick can be viewed as an atomic approximation for exchange and correlation [14] in which exchange and correlation energies cancel the ionic and Hartree potentials outside the atomic cell.

The Wigner-Seitz trick can be extended to a more general approximation for exchange and correlation, which is based on atomic exchange and correlation rather than that of the homogeneous electron gas. The construction starts with the best atomic potential available which can be obtained either from first principles or from observed atomic levels. The electronic potential at each point in the solid is just the atomic potential of greatest magnitude, usually the potential of the nearest atom. Just as for the Wigner-Seitz trick, the approximation for exchange and correlation implied by this trick is that the exchange and correlation potential of each atom cancels its ionic and Hartree potentials outside its atomic cell.

The main difference between Wigner-Seitz exchange and correlation (WSA) and the LDA is that the former is non-local and the latter is local. The WSA exchange and correlation energy at each point in the solid does not depend only on the electron density at that point, but rather on the dominant nearby atom. Applications of WSA to MgO [14] suggest that it gets much better band gaps than LDA, and its close relation to the Mattheiss prescription [19] supports this.

It is possible to go beyond WAS by taking into account the polarizability of atoms as was done by Annett and Haydock [20] in calculations of interactions between helium atoms and metal surfaces. As electrons move inside one atom, they polarize the surrounding solid and thereby lower their energies; this effect leads to Van der Waals forces when the atoms are far apart. It is not included in the WAS exchange and correlation energy, but the effect in the homogeneous

electron gas is included in LDA.

ACKNOWLEDGMENTS

REFERENCES

1.   Roger Haydock in Mathematics and its Applications vol. 99, edited by Golden, Grimmett, James, Milton, and Sen (Springer, Berlin, 1997), 107-114.

2.   P. Hohenberg and W. Kohn, Phys. Rev. **136,** 864 (1964).

3.   W. Kohn and L.J. Sham, Phys. Rev. **140,** 1133 (1965).

4.   J. Harris, Phys. Rev. B **31,** 1770 (1985).

5.   W.M.C. Foulkes and Roger Haydock, Phys. Rev. B **39,** 12520 (1989).

6.   H. Wendel and R.M. Martin, Phys. Rev. B **19,** 5251 (1979).

7.   K.W. Jacobsen, J.K. Norskov, and M.J. Puska, Phys. Rev. B **35,** 7423 (1987).

8.   Roger Haydock and Volker Heine, Comments on Condensed Matter Physics **18,** 83 (1997).

9.   A. Gibson, Roger Haydock, and J.P. LaFemina, Phys. Rev. B. **50,** 2582 (1994).

10.   S. Blumberg and Roger Haydock, Phys. Rev. B **45,** 1550 (1992).

11.   J.C. Slater and G.F. Koster, Phys. Rev. **94,** 1498 (1954).

12.   A. Gibson and Roger Haydock, Phys. Rev. B. **47,** 9229 (1993).

13.   Roger Haydock in Solid State Physics vol. 35, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, London and New York, 1980), 215-294.

14.   A. Gibson, Roger Haydock, and J.P. LaFemina, J. Vac. Sci. Technol. A **10,** 2361 (1992).

15.   W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes (Cambridge University Press, Cambridge, 1986).

16.   J. Reitz in Solid State Physics vol. 1, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic Press, London and New York, 1955).

17.   F. Seitz, The Modern Theory of Solids (McGraw-Hill, New York, 1940).

18.   J.E. Inglesfield, Mol. Phys. **37,** 3 and 889 (1979).

19.   L.F. Mattheiss, Phys. Rev. **133,** A1399 (1964).

20.   J.F. Annett, and Roger Haydock, Phys. Rev. B **34,** 6860 (1986).

A LCAO-OO APPROACH TO THE CALCULATION OF ELECTRONIC PROPERTIES OF MATERIALS

P. POU, R. PEREZ, J. ORTEGA, F. FLORES
Física Teórica de la Materia Condensada, C-V, Universidad Autónoma de Madrid, 28049 Madrid, Spain.

ABSTRACT

We present a selfconsistent LCAO approach for describing the electronic properties of materials. This method introduces many-body effects by means of a new approach, whereby a local orbital potential is defined by calculating the exchange-correlation energy as a function of the different orbital occupancies. A LCAO-pseudopotential is also introduced, keeping all the calculations within the context of the local orbital basis. We have applied the method to the calculation of simple molecules and crystals, in both cases we find results that confirm the validity of our approach.

INTRODUCTION

Plane waves based electronic structure methods are well-established worldwide [1] and are regularly used by a variety of people, many of whom are not experts in electronic structure methods. These methods can be efficiently applied only to systems if up to a few hundred atoms, using the latest parallel computers, and there is a need to implement accurate methods to very large systems with complex physical conditions. We believe this goal can be best accomplished by means of an appropriate combination of density-functional methods and a local-orbital basis set [2-4]. The use of a local-orbital basis set is conceptually straightforward and computationally very efficient since it allows one to take full advantage of algorithms that scale linearly with system size and get optimal speed up of computer implementation [5]. These are attributes which have been shown to be highly desirable when solving molecular dynamics problems where computational speed is an important factor, such as calculations involving complex structures [6].

In this paper we present a selfconsistent LCAO (Lineal Combination of Atomic Orbitals) method, where the electronic properties of materials are described using a localized basis set associated with the atoms forming the system. This method introduces many-body effects by means of a new approach whereby a local orbital potential is defined by calculating the exchange-correlation energy as a function of the different orbital occupancies.

Although some of the techniques used in this approach have been described in previous papers [7-9], this is the first time we present a complete description of the LCAO-OO (Orbital Occupancy) method that allows us to calculate the electronic properties using a fully local basis approach.

The following sections cover the following steps:
a) Firstly, we present our basis set and discuss a LCAO-pseudopotential that incorporates all the effects that the core electrons have on the valence bands.
b) In a second step, we present the approximate hamiltonian that will allow us to get an accurate electronic description of the materials.
c) Then, we analyze the many-body properties of the valence electrons and discuss our Orbital Occupancy approach by introducing the orbital occupancies as the

independent variables of the many-body problem. We shall show how to reduce the many-body effects to an effective one-electron hamiltonian, and discuss how to obtain the total energy.

d) Finally, we present some results we have obtained for simple molecules and the Si bulk crystal. For Si, the band-structure, tight-binding parameters, and the total energy as a function of the crystal lattice parameter are presented and discussed.

GENERAL HAMILTONIAN AND A LCAO-PSEUDOPOTENTIAL.

In our LCAO approach the starting point is defined by the wavefunction basis, $\psi_\alpha(\vec{r} - \vec{R}_i)$, where α refers to the atomic orbital of the i-site. We write $(i, \alpha) = \nu$ and introduce the Löwdin's orthogonal basis:

$$\phi_\mu = \sum_\nu (S^{-1/2})_{\mu\nu} \psi_\nu \tag{1}$$

where $S_{\mu\nu} = <\psi_\mu | \psi_\nu>$, is the overlap between orbitals μ and ν.

The initial basis, $\psi_\alpha(i)$, might be defined by the atomic wavefunction of the atom i but, in general, we use contracted atomic wavefunctions that we calculate for the isolated atom by introducing an external potential of the form $a(r/R_0)^n$. Typically we take, following other authors [6,10], $a = 1$ Hartree $n = 6$ and change R_0 in order to improve the basis. The role of the external potential, $a(r/R_0)^n$, is to simulate the effect of the environment of the constituents of the material on the wavefunctions. In these calculations, the atomic wavefunctions are a lineal combination of Slater-type functions, which define the initial basis set, $\psi_\alpha(i)$, of our problem.

The electronic properties of our system are defined by the general many-body hamiltonian:

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_{i,k} \frac{Z_k}{\left|\vec{r}_i - \vec{R}_k\right|} + \frac{1}{2}\sum_{i,j\neq i}\frac{1}{\left|\vec{r}_i - \vec{r}_k\right|} \tag{2}$$

that can be rewritten in a second quantization language as follows:

$$\hat{H} = \sum_{\nu,\sigma} \varepsilon_\nu \hat{n}_{\nu\sigma} + \sum_{\mu\neq\nu,\sigma} t_{\mu\nu} c_{\mu\sigma}^+ c_{\nu\sigma} + \sum_{\substack{\mu\nu\lambda\omega \\ \sigma\sigma'}} O_{\mu\nu\lambda\omega} c_{\mu\sigma}^+ c_{\nu\sigma}^+ \cdot c_{\lambda\sigma} \cdot c_{\omega\sigma} \tag{3}$$

where we have used the Löwdin's basis set, ϕ_μ. In equation (3):

$$\varepsilon_\nu = \int \phi_\nu(\vec{r})(-\frac{1}{2}\nabla^2 - \sum_k \frac{Z_k}{\left|\vec{r} - \vec{R}_k\right|})\phi_\nu(\vec{r})d^3r,$$

$$t_{\nu\mu} = \int \phi_\nu(\vec{r})(-\frac{1}{2}\nabla^2 - \sum_k \frac{Z_k}{\left|\vec{r} - \vec{R}_k\right|})\phi_\mu(\vec{r})d^3r, \tag{4}$$

$$O_{\mu\nu\lambda\omega} = \int \phi_\mu(\vec{r})\phi_\nu(\vec{r}')\frac{1}{\left|\vec{r} - \vec{r}'\right|}\phi_\lambda(\vec{r}')\phi_\omega(\vec{r})d^3rd^3r'$$

We stress that eqns. (3) and (4) include both the valence and core electrons. Our first task is to simplify hamiltonian (3) by introducing an effective pseudopotencial that allows us to

eliminate the core electrons. We can calculate that pseudopotencial by taking advantage of the small overlap, S_{vc}, between the wavefunctions of the valence (v) and core (c). In this limit, we can calculate the different parameters of hamiltonian (3), eqns. (4), up to second order in S_{vc} and obtain the effect of the core on the valence electrons [7]. Here, for the sake of brevity, we present a simplified argument by considering only the interaction of a single v-valence orbital of atom 1 with another individual c-level of atom 2 (see fig.1), and neglect all the many-body interactions. Then, hamiltonian (3) is reduced to:

$$\hat{H} = \sum_{\sigma} \left[\varepsilon_v \hat{n}_{v\sigma} + \varepsilon_c \hat{n}_{c\sigma} + T_{vc}(c_{v\sigma}^+ c_{c\sigma} + c_{c\sigma}^+ c_{v\sigma}) \right] \tag{5a}$$

where

$$\varepsilon_v = \int \phi_v(\vec{r}) \left[-\frac{1}{2}\nabla^2 + V_1(\vec{r}) + V_2(\vec{r}) \right] \phi_v(\vec{r}) d^3r,$$

$$\varepsilon_c = \int \phi_c(\vec{r}) \left[-\frac{1}{2}\nabla^2 + V_1(\vec{r}) + V_2(\vec{r}) \right] \phi_c(\vec{r}) d^3r, \tag{5b}$$

$$T_{vc} = \int \phi_v(\vec{r}) \left[-\frac{1}{2}\nabla^2 + V_1(\vec{r}) + V_2(\vec{r}) \right] \phi_c(\vec{r}) d^3r$$

Here, potentials $V_1(\vec{r})$ and $V_2(\vec{r})$ simulate, in this simplified discussion, the total atomic potentials of the atoms 1 and 2. Notice also that the Löwdin's wavefunctions, ϕ_v and ϕ_c, are related to the atomic ones, ψ_v and ψ_c, by eqn. (1). By using this eqn, and expanding $S^{-1/2}$ up to second order in the overlap S_{vc}, we get the following eqns.:

$$\varepsilon_v = E_v^{(0)} + \frac{1}{4}S_{vc}^2(E_v^{(0)} - E_c^{(0)}) - S_{vc}T_{vc} \tag{6a}$$

$$\varepsilon_c = E_c^{(0)} - \frac{1}{4}S_{vc}^2(E_v^{(0)} - E_c^{(0)}) - S_{vc}T_{vc} \tag{6b}$$

where $E_v^{(0)}$ and $E_c^{(0)}$ are the energy levels for the isolated atoms.

In a further step, we can show that for a localized core level, T_{vc} can be accurately approximated by [7]:

$$T_{vc} = -\frac{1}{2}S_{vc}(E_v^{(0)} - E_c^{(0)}) \tag{7}$$
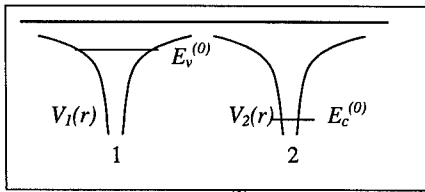
Fig. 1: Valence level $E_v^{(0)}$ and core level $E_c^{(0)}$.

47

From eqns. (6) and (7) we can obtain the effect of the valence-core interactions on the atomic levels. Considering that the effect of T_{vc} is to shift both levels by the quantity $\pm T_{vc}^2 / (E_v^{(0)} - E_c^{(0)})$, we can write for the valence and core level shifts:

$$\delta\varepsilon_v = \varepsilon_v - E_v^{(0)} = \frac{1}{4} S_{vc}^2 (E_v^{(0)} - E_c^{(0)}) - S_{vc} T_{vc} + \frac{T_{vc}^2}{(E_v^{(0)} - E_c^{(0)})} =$$

$$= S_{vc}^2 (E_v^{(0)} - E_c^{(0)}) \tag{8}$$

$$\delta\varepsilon_c = \varepsilon_c - E_c^{(0)} = -\frac{1}{4} S_{vc}^2 (E_v^{(0)} - E_c^{(0)}) - S_{vc} T_{vc} - \frac{T_{vc}^2}{(E_v^{(0)} - E_c^{(0)})} = 0$$

This is our main result for the level shifts as a function of the one-electron interaction between the valence and core electrons. Eqns. (8) shows that the valence level is shifted by $S_{vc}^2 (E_v^{(0)} - E_c^{(0)})$, while the core level remains unmodified. This result can be easily generalized to several core levels; then the valence level shift is given by:

$$\delta\varepsilon_v = \sum_c S_{vc}^2 (E_v^{(0)} - E_c^{(0)}) \tag{9}$$

A similar argument applied to the valence-valence hopping integral, say $T_{vv'}$, also yields a hopping shift, $\delta T_{vv'}$, due to the valence-core overlap that can be written as follows:

$$\delta T_{vv'} = \sum_c S_{vc} S_{cv'} \left(\frac{E_v^{(0)} + E_{v'}^{(0)}}{2} - E_c^{(0)} \right) \tag{10}$$

Eqns. (9) and (10) provide the pseudopotential matrix elements, in the LCAO basis. Although v and v' have been assumed to be atomic orbitals, eqns. (9) and (10) are valid whatever the basis set we use. (Notice that with very good accuracy we can replace $E_v^{(0)}$ and $E_{v'}^{(0)}$ by the energy, E, of the valence wavefunctions we are interested in).

Many-body contributions to the pseudopotential can also be calculated by starting from the general hamiltonian (3). Without giving any details, let us only mention that the general results we obtain for δE_v and $\delta T_{vv'}$ are the following:

$$\delta E_v = \sum_c S_{vc} S_{cv} (E - E_c^{(0)}) + \sum_c J_{vc}^{(0)} n_c + \sum_c (J_{vc}^{(0)} S_{vc}^2 - J_{vc}^{x(0)}) n_{c\sigma}$$

$$\delta T_{vv'} = \sum_c S_{vc} S_{cv'} (E - E_c^{(0)}) + \sum_c h_{c,vc}^{(0)} n_c + \sum_c \left(\frac{J_{vc}^{(0)} + J_{v'c}^{(0)}}{2} S_{vc} S_{cv'} - h_{c,vc}^{x(0)} \right) n_{c\sigma} \tag{11;}$$

$$J_{vc}^{(0)} = \int \psi_v^2(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \psi_c^2(\vec{r}') d^3r d^3r'$$

$$J_{vc}^{x(0)} = \int \psi_v(\vec{r}) \psi_c(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \psi_v(\vec{r}') \psi_c(\vec{r}') d^3r d^3r'$$

$$h_{c,vv'}^{(0)} = \int \psi_c^2(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \psi_v(\vec{r}') \psi_{v'}(\vec{r}') d^3r d^3r' \tag{12}$$

$$h_{c,vv'}^{x(0)} = \int \psi_c(\vec{r}) \psi_v(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \psi_c(\vec{r}') \psi_{v'}(\vec{r}') d^3r d^3r'$$

In eqn. (11), $n_c = n_{c\sigma} + n_{c\bar{\sigma}}$, $n_{c\sigma}$ is the occupation number of the $(c\sigma)$-orbital. Notice that $J_{vc}^{(0)} n_c$ and $h_{c,vv'}^{(0)} n_c$ define the coulomb corrections of the bare core-levels, while the last terms in each eqn. yield an exchange correction to the potential. In the following we shall define the LCAO-pseudopotential by the eqns.:

$$V_{vv}^{ps} = \delta E_v$$
$$V_{vv'}^{ps} = \delta T_{vv'}$$

(13)

APPROXIMATE VALENCE HAMILTONIAN

Having introduced the pseudopotencial, $V_{\mu v}^{ps}$, we can write down the following hamiltonian for the valence electrons:

$$\hat{H} = \sum_{v,\sigma} (\varepsilon_v + V_{vv}^{ps}) \hat{n}_{v\sigma} + \sum_{\mu \neq v,\sigma} (t_{\mu v} + V_{\mu v}^{ps}) c_{\mu\sigma}^+ c_{v\sigma} + \sum_{\substack{\mu v \lambda \omega \\ \sigma\sigma'}} O_{\mu v \lambda \omega} c_{\mu\sigma}^+ c_{v\sigma'}^+ c_{\lambda\sigma'} c_{\omega\sigma}$$

(14)

Now, we simplify this eqn. by reducing it to the following hamiltonian:

$$\hat{H} = \sum_{v,\sigma} (\varepsilon_v + V_{vv}^{ps}) \hat{n}_{v\sigma} + \sum_{\mu \neq v,\sigma} (t_{\mu v} + V_{\mu v}^{ps} + \sum_\lambda h_{\lambda,\mu v} \hat{n}_\lambda - \sum_\lambda h_{\lambda,\mu v}^x \hat{n}_{\lambda\sigma}) c_{\mu\sigma}^+ c_{v\sigma} +$$
$$+ \sum_v U_v \hat{n}_{v\uparrow} \hat{n}_{v\downarrow} + \frac{1}{2} \sum_{\mu \neq v,\sigma} (J_{\mu v} \hat{n}_\mu \hat{n}_v - J_{\mu v}^x \hat{n}_{\mu\sigma} \hat{n}_{v\sigma})$$

(15);

$$U_v = \int \phi_v^2(\vec{r}) \frac{1}{|\vec{r} - \vec{r'}|} \phi_v^2(\vec{r'}) d^3r d^3r'$$

$$J_{\mu v} = \int \phi_\mu^2(\vec{r}) \frac{1}{|\vec{r} - \vec{r'}|} \phi_v^2(\vec{r'}) d^3r d^3r'$$

$$J_{\mu v}^x = \int \phi_\mu(\vec{r}) \phi_v(\vec{r}) \frac{1}{|\vec{r} - \vec{r'}|} \phi_\mu(\vec{r'}) \phi_v(\vec{r'}) d^3r d^3r'$$

(16)

$$h_{\lambda,\mu v} = \int \phi_\lambda^2(\vec{r}) \frac{1}{|\vec{r} - \vec{r'}|} \phi_\mu(\vec{r'}) \phi_v(\vec{r'}) d^3r d^3r'$$

$$h_{\lambda,\mu v}^x = \int \phi_\lambda(\vec{r}) \phi_\mu(\vec{r}) \frac{1}{|\vec{r} - \vec{r'}|} \phi_\lambda(\vec{r'}) \phi_v(\vec{r'}) d^3r d^3r'$$

where Löwdin´s wavefunctions, ϕ_μ, have been used (compare with eqns. (12) which use atomic rather than Löwdin´s wavefunctions).

In ref. [11], Bolcatto *et al* have shown that eqn. (15) provides a very accurate solution for the bonding energy of diatomic molecules.

It is convenient to analyze the physical meaning of hamiltonian (15) by introducing the following hartree-like approximation:

$$\hat{H}_0 = \sum_{v,\sigma} (\varepsilon_v + V_{vv}^{ps} + U_v n_{v\bar{\sigma}} + \sum_\mu J_{\mu v} n_\mu - \sum_\mu J_{\mu v}^x n_{\mu\sigma}) \hat{n}_{v\sigma} +$$
$$+ \sum_{\mu \neq v,\sigma} (t_{\mu v} + V_{\mu v}^{ps} + \sum_\lambda h_{\lambda,\mu v} n_\lambda - \sum_\lambda h_{\lambda,\mu v}^x n_{\lambda\sigma}) c_{\mu\sigma}^+ c_{v\sigma}$$

(17)

where $n_\mu = n_{\mu\sigma} + n_{\mu\bar\sigma}$, $n_{\mu\sigma} = \langle \hat{n}_{\mu\sigma} \rangle$ is the mean value of the operator $\hat{n}_{\mu\sigma}$. Eqn. (17) is the second quantization representation of the following one-electron hamiltonian:

$$\hat{H}_0 = -\frac{1}{2}\nabla^2 - \sum_k \frac{Z_k}{|\vec{r} - \vec{R}_k|} + V^{ps}(\vec{r}) + \sum_v n_v \int \frac{\phi_v^2(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3 r' - \sum_{v\sigma} n_{v\sigma} \frac{|\phi_v(\vec{r})\rangle\langle\phi_v(\vec{r}')|}{|\vec{r} - \vec{r}'|} =$$

$$= \hat{H}_{0,h} - \sum_{v\sigma} n_{v\sigma} \frac{|\phi_v(\vec{r})\rangle\langle\phi_v(\vec{r}')|}{|\vec{r} - \vec{r}'|}$$
(18)

Eqn. (18) shows that in our approximation, the hartree-terms are defined by the potential created by the atomic charges, $n_{v\sigma}\phi_v^2(\vec{r})$, and the attractive potential of the nuclei, $-Z/|\vec{r} - \vec{R}|$. This is similar to the atomic-like hartree potential usually introduced in more conventional LDA-methods based on the Harris' functional [2]. A non-local potential, associated with the term $n_{v\sigma}|\phi_v(\vec{r})\rangle\langle\phi_v(\vec{r}')|$ and describing an exchange interaction, also appears in eqn. (18).

In the next section we shall discuss how to include other exchange and correlation effects in eqn. (17).

MANY-BODY TERMS

We start our discussion of the many-body terms of hamiltonian (15) by writing down the following contribution:

$$\hat{H}^{m.b.} = +\sum_v U_v \hat{n}_{v\uparrow}\hat{n}_{v\downarrow} + \frac{1}{2}\sum_{\mu \neq v,\sigma}(J_{\mu v}\hat{n}_\mu\hat{n}_v - J_{\mu v}^x\hat{n}_{\mu\sigma}\hat{n}_{v\sigma})$$
(19)

As discussed in ref. [9], we can introduce the orbital occupation number, $n_{\mu\sigma}$, as the independent variable that plays the same role $n(\vec{r})$ does in conventional LD approaches. This means that we can define the following function, $E^{m.b.}[n_{1\sigma}, n_{2\sigma}, ...]$ associated with hamiltonian (19), playing the same role that the many-body potential, $E^{m.b.}[n(\vec{r})]$, does in LDA. In particular, knowing $E^{m.b.}[n_{\mu\sigma}]$ allows us to introduce the following local potential:

$$V_{\mu\sigma}^{m.b.} = \frac{\partial E^{m.b.}[n_{\mu\sigma}]}{\partial n_{\mu\sigma}}$$
(20)

Then, hamiltonian (19) can be replaced by the following term:

$$\sum_{\mu\sigma} V_{\mu\sigma}^{m.b.}\hat{n}_{\mu\sigma}$$
(21)

that plays the role of a local many-body potential in our OO-approach.

Still, we have to calculate $E^{m.b.}[n_{\mu\sigma}]$. We proceed in the usual way, writing:

$$E^{m.b.}[n_{\mu\sigma}] = E^H[n_{\mu\sigma}] + E^{XC}[n_{\mu\sigma}]$$
(22)

where the hartree energy associated with eqn. (19) is given by:

$$E^H[n_{\mu\sigma}] = \sum_\mu U_\mu n_{\mu\uparrow} n_{\mu\downarrow} + \frac{1}{2} \sum_{\mu\neq\nu,\sigma} (J_{\mu\nu} n_\mu n_\nu - J^x_{\mu\nu} n_{\mu\sigma} n_{\nu\sigma}) \tag{23}$$

and

$$V^H_{\mu\sigma} = U_\mu n_{\mu\bar\sigma} + \sum_\nu (J_{\mu\nu} n_\nu - J^x_{\mu\nu} n_{\nu\sigma}) \tag{24}$$

In ref. [9], we have shown that the exchange energy can be written as follows:

$$E^x[n_{\mu\sigma}] = -\frac{1}{2} \sum_{\mu\neq\nu,\sigma} (J_{\mu\nu} - J^x_{\mu\nu}) \langle c^+_{\mu\sigma} c_{\nu\sigma} \rangle \langle c^+_{\nu\sigma} c_{\mu\sigma} \rangle = -\frac{1}{2} \sum_{\mu\neq\nu,\sigma} J_\mu n_{\mu\sigma} (1 - n_{\mu\sigma}) \tag{25}$$

where the mean electron charge, $n_{\mu\sigma}$, filling the $\mu\sigma$-level, interacts with its hole $(1 - n_{\mu\sigma})$ via the Coulomb interaction, J_μ. Typically, J_μ is the Coulomb interaction between electrons located in nearest neighbor sites; this means that in most cases, the electron hole is practically localized in nearest neighbor atoms.

Correlation effects are mainly associated with transferring part of the exchange hole to the same atom of the $\mu\sigma$-orbital. If a fraction, say f_μ ($f_\mu < 1$), of this hole is located in this atom, the correlation energy should be given by [9]:

$$E^c[n_{\mu\sigma}] = -\frac{1}{2} \sum_{\mu\neq\nu,\sigma} f_\mu U_\mu n_{\mu\sigma} (1 - n_{\mu\sigma}) \tag{26}$$

where U_μ is an average of the intrasite Coulomb interactions between different orbitals. At the same time, the exchange energy should be reduced by the factor $(1 - f_\mu)$ due to having only a fraction of the hole $(1 - f_\mu)(1 - n_{\mu\sigma})$, outside the atom. Then, the total exchange-correlation energy is given by:

$$E^{XC}[n_{\mu\sigma}] = -\frac{1}{2} \sum_{\mu\neq\nu,\sigma} (1 - f_\mu) J_\mu n_{\mu\sigma} (1 - n_{\mu\sigma}) - \frac{1}{2} \sum_{\mu\neq\nu,\sigma} f_\mu U_\mu n_{\mu\sigma} (1 - n_{\mu\sigma}) =$$
$$= -\frac{1}{2} \sum_{\mu\neq\nu,\sigma} J_\mu n_{\mu\sigma} (1 - n_{\mu\sigma}) - \frac{1}{2} \sum_{\mu\neq\nu,\sigma} f_\mu (U_\mu - J_\mu) n_{\mu\sigma} (1 - n_{\mu\sigma}) \tag{27}$$

and the corresponding potential by:

$$V^{XC}_{\mu\sigma} = -J_\mu (\frac{1}{2} - n_{\mu\sigma}) - f_\mu (U_\mu - J_\mu)(\frac{1}{2} - n_{\mu\sigma}) \tag{28}$$

Then the many-body potential is defined by eqns. (24) and (28).

We should comment that in our approximation the selfinteraction correction is automatically included (the interaction between electrons of the same spin occupying the same orbital is taken zero). We also note that the Hartree-Fock approximation for the total energy of hamiltonian (19) is equivalent to considering only in our case the exchange energy given by eqn. (25). Moreover, going beyond the LD approximation can be achieved by substituting the correlation energy, eqn. (26), for a selfenergy [15] that describes the intraatomic correlation fluctuations. This is similar to the GW approximation introduced in more conventional LD calculations.

METHOD OF CALCULATION

In this section we present a summary of the method used to calculate the electronic properties of materials.

Firstly, hamiltonian (15) is further approximated by:

$$\hat{H} = \sum_{v,\sigma}(\varepsilon_v + V_{vv}^{ps})\hat{n}_{v\sigma} + \sum_{\mu \neq v,\sigma}(t_{\mu v} + V_{\mu v}^{ps} + \sum_\lambda h_{\lambda,\mu v}n_\lambda - \sum_\lambda h_{\lambda,\mu v}^x n_{\lambda\sigma})c_{\mu\sigma}^+ c_{v\sigma} +$$

$$+ \sum_v U_v \hat{n}_{v\uparrow}\hat{n}_{v\downarrow} + \frac{1}{2}\sum_{\mu \neq v,\sigma}(J_{\mu v}\hat{n}_\mu \hat{n}_v - J_{\mu v}^x \hat{n}_{\mu\sigma}\hat{n}_{v\sigma})$$

(29)

where $h_{\lambda,\mu v}\hat{n}_\lambda$ and $h_{\lambda,\mu v}^x \hat{n}_\lambda$ have been replaced by $h_{\lambda,\mu v}n_\lambda$ and $h_{\lambda,\mu v}^x n_\lambda$, respectively. We have checked this is a reasonable approximation to hamiltonian (15), because the exchange contributions of both terms tend to cancel each other (notice that $h_{\mu,\mu v} = h_{\mu,\mu v}^x$). In any case, it is straightforward to obtain perturbatively the contribution coming from the exchange terms associated with the interactions $h_{\lambda,\mu v}$ and $h_{\lambda,\mu v}^x$, by calculating:

$$-\sum_{\lambda,\mu v,\sigma}(h_{\lambda,\mu v} - h_{\lambda,\mu v}^x)\langle c_{\lambda\sigma}^+ c_{\mu\sigma}\rangle \langle c_{v\sigma}^+ c_{\lambda\sigma}\rangle$$

(30)

where $\langle c_{\lambda\sigma}^+ c_{\mu\sigma}\rangle$ is the expectation value of $(c_{\lambda\sigma}^+ c_{\mu\sigma})$ in the ground state of the effective hamiltonian (see eqn. (31) below).

Eqn. (29) is analyzed by using our OO-approach. This amounts to introducing the following effective hamiltonian:

$$\hat{H}^{eff} = \sum_{v,\sigma}(\varepsilon_v + V_{vv}^{ps} + V_v^{m.h.})\hat{n}_{v\sigma} + \sum_{\mu \neq v,\sigma}(t_{\mu v} + V_{\mu v}^{ps} + \sum_\lambda h_{\lambda,\mu v}n_\lambda - \sum_\lambda h_{\lambda,\mu v}^x n_{\lambda\sigma})c_{\mu\sigma}^+ c_{v\sigma}$$ (31)

This one-electron hamiltonian is solved and the ground state, $|\Phi_0\rangle$, and occupation numbers, $n_{\mu\sigma}$, are obtained selfconsistently. Then, the total energy is calculated, as in conventional LD-methods, by adding the one-electron and the many-body energies:

$$E_{TOT} = \sum_n e_n - E_{DC} + E_{ion-ion} + E^{XC} - \sum_{v\sigma}V_{v\sigma}^{XC}n_{v\sigma}$$

(32)

The first term of eqn. (32) is the sum of occupied eigenvalues of the effective hamiltonian, $\hat{H}^{eff}$, E_{DC} is the double-counting correction for the hartree-like contributions (eqn. (18)), and $E_{ion-ion}$ is the repulsive energy between nuclear charges, Z_k.

Finally, a few words should be said about the numerical methods used to calculate the different quantities of hamiltonian (15) and (19). The hartree terms, $\hat{H}_{0,h}$, defined by one-electron hamiltonian (18) are calculated in the atomic basis, ψ_v, and transformed to the Löwdin basis by the eqn.:

$$\langle \phi_\mu | \hat{H}_{0,h} | \phi_v \rangle = \sum_{\mu'v'}(S^{-\frac{1}{2}})_{\mu\mu'}\langle \psi_{\mu'} | \hat{H}_{0,h} | \psi_{v'} \rangle (S^{-\frac{1}{2}})_{v'v}$$

(33)

We have found practical to replace the Löwdin wavefunctions by the atomic ones in the calculation of the hartree potential [4]:

52

$$\sum_{v\sigma} n_{v\sigma} \int \frac{\phi_v^2(\vec{r}')}{\left|\vec{r}-\vec{r}'\right|} d^3 r' \approx \sum_{v\sigma} n_{v\sigma} \int \frac{\psi_v^2(\vec{r}')}{\left|\vec{r}-\vec{r}'\right|} d^3 r' \tag{34}$$

This is an accurate procedure, once E_{DC} is appropriately calculated taking into account eqn. (34) (details will be published elsewhere).

We still have to calculate U_μ, $J^x_{\mu\nu}$, $J_{\mu\nu}$, $h_{\lambda,\mu\nu}$ and $h^x_{\lambda,\mu\nu}$ in order to obtain the exchange-correlation corrections to the hartree potential. We have calculated those parameters (see eqns. (15)), by obtaining, in a first step, the Löwdin's wavefunction, ϕ_μ, which, in the cases presented below, extend practically to the fourth nearest neighbors. By using a small cluster around sites μ and /or ν, we have calculated the integrals of eqns. (17) for U_μ, $J^x_{\mu\nu}$, $J_{\mu\nu}$, $h_{\lambda,\mu\nu}$ and $h^x_{\lambda,\mu\nu}$ up to the nearest neighbors that yield relevant contributions to the exchange correlation energy.

RESULTS AND DISCUSSION

As an application of the LCAO-OO approach, we present calculations of a simple molecule, CH, and Si bulk crystal.

In fig. 2, we show the bond energy of CH as a function of inter-atomic distance. In this calculation we have used the following basis:
for H: 1s orbital,
for C: 1s, 2s and 2p orbitals.

In this calculation we have used the double-zeta functions given in tables compiled by Clementi and Roetti [12].

In fig. 2, we also show the complete many-body solution of eqn. (3). This solution has been obtained calculating all the different parameters of hamiltonian (3) by means of eqn. (4). Then, the ground state wavefunction of this hamiltonian is calculated within the configuration space associated with all the valence electron redistributed among the 2s and 2p orbitals of C and the 1s-orbital of H.
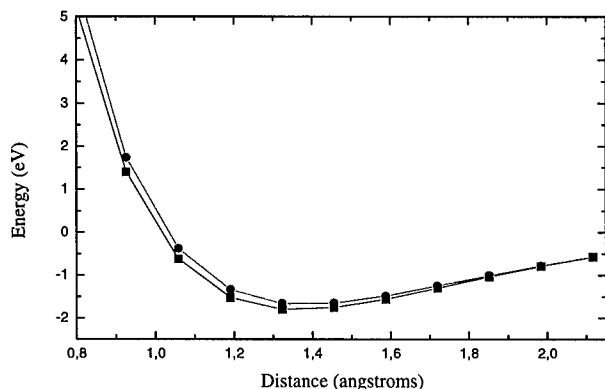
Fig. 2: Bond energy of CH as function of inter-atomic distance. (■) Many-body solution of eqn. (3). (●) Our calculation.

The comparison between the "exact" solution and the one given by our method is excellent. Both binding energies are very similar and the equilibrium distances are almost identical. Notice that the binding energy of the initial hamiltonian is ~ 1.8 eV, lower than the

experimental value of 3.6 eV. The difference is obviously related to the minimal basis that we are using to analyze the molecule.

Bulk Si has been analyzed using the following minimal basis:

Si: 1s, 2s, 2p, 3s and 3p orbitals.

In this calculation we have used the Slater-like single-zeta functions. In this crystal, we have found that the effect of the orbital relaxation is small, and the results presented below were calculated with a slightly contracted single-zeta function w.r.t. the Clementi and Roetti's wavefunctions [12].

Fig. 3 shows our calculation of the cohesive energy of Si per atom as a function of the crystal lattice parameter. This curve yields a cohesive energy of $\sim$ 2.3 eV, a crystal lattice parameter of $\sim$ 5.5 Å, and a bulk modulus, $-V\partial^2 E/\partial V^2$, of $\sim$ 72 GPa . The crystal lattice parameter and the bulk modulus are in reasonable agreement with the experimental evidence, although the cohesive energy is too small by $\sim$ 2 eV. This difference can be attributed to the minimal basis we have used for Si; other calculations [13] only obtain a good convergence for the cohesive energy when 3d orbitals are included in the Si-basis. From independent H-F calculations [14], we deduce that the correlation energy per atom is around 1.5 eV. Since, in our calculation this correlation energy per atom is only $\sim$ 0.8 eV, we conclude that we still need $\sim$ 0.7 eV of correlation energy (due to a larger configurational space), and $\sim$ 1.3 eV associated with the use of a minimal basis.
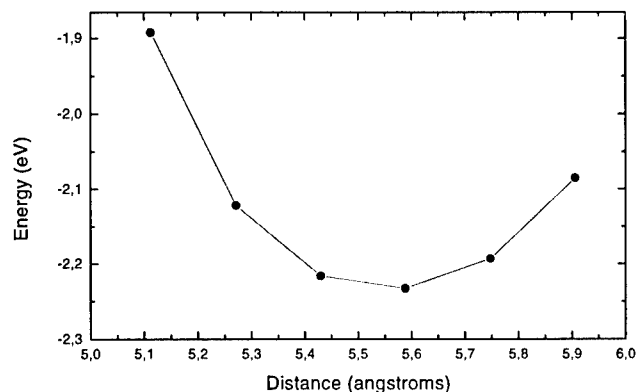
Fig. 3: Cohesive energy per atom of Si bulk crystal as a function of the crystal parameter distance.

Fig. 4 shows our result for the electronic band structure of Si. The valence band shows a general agreement with other calculations [13,14], although the conduction band does not seem to be completely converged. This is probably also due to not having included the Si 3d orbitals. We would expect them to substantially modify the conduction band.

CONCLUSIONS

In conclusion, the results discussed for molecule (CH) and crystal (Si) show good agreement with the general solution of these systems and allow us to assess the validity and accuracy of the method presented in this paper. We stress that this method is the first fully selfconsistent approach that describes all the properties of the system within the limits of a strict

LCAO method. We believe that this could represent an advantage for the preparation of fast and efficient computational codes. Work in this direction is currently under investigation in our labs.
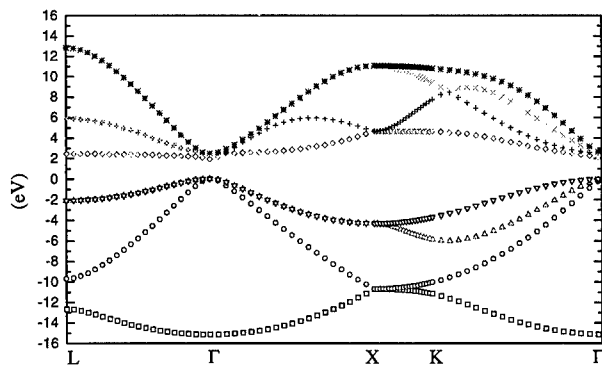
Fig. 4: Electronic band structure of Si bulk crystal.

ACKNOWLEDGMENTS

REFERENCES

1. R. Car and M. Parrinello, Phys. Rev. Lett. **55**,2471 (1985); M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
2. J. Harris, Phys. Rev. B **31**, 1770 (1985).
3. O. F. Sankey and D. J. Niklewski, Phys. Rev, B **40**, 3979 (1989); Z. Lin and J. Harris, J. Phys. Condens. Matter **5**, 1055 (1993); D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert and R. Kaschuer, Phys. Rev. B **51**, 12947 (1995); P. Ordejón, E. Artacho and J. M. Soler, Phys. Rev. B **53**, 10441 (1996).
4. A. A. Demkov, J. Ortega, O. F. Sankey and M. P. Grumbach, Phys. Rev. B **52**, 1618 (1995).
5. See for example, F. Mauri *et al*, Phys. Rev. B **47**, 9973 (1993); P. Ordejón, D. A. Drabold, M. P. Grumbach and R. M. Martin, Phys. Rev. B **48**, 14646 (1993).
6. A. P. Horsefield, Phys. Rev. B **56**, 6594 (1996); S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994).
7. E. C. Goldberg, A. Martin-Rodero, R. Monreal and F. Flores, Phys. Rev. B **39**, 5684 (1988).
8. F. J. Garcia-Vidal, A. Martin-Rodero, F. Flores, J. Ortega and R. Pérez, Phys. Rev. B **44**, 11412 (1991).
9. F. J. Garcia-Vidal, J. Merino, R. Pérez, R. Rincón, J. Ortega and F. Flores, Phys. Rev. B **50**, 10537 (1994).
10. H. Eschrig and I. Bergert, Phys. Stat. Sol. B **90**, 621 (1978).
11. P. G. Bolcatto, E. C. Goldberg and C. G. Passoppi, Phys. Rev. A **50**, 4643 (1994).
12. Clementi and Roetti, Atomic and Nuclear Data Tables, **14**, 177, (1974).
13. R. W. Jansen and O. F. Sankey, Phys. Rev. B **36**, 6520 (1987).

14. R. Dovesi, M. Causà and G. Angonoa, Phys. Rev. B **24**, 4177 (1981); M. Causà, R. Dovesi and C. Roetti, Phys. Rev. B **43**, 11937 (1991).

15. H. Kajueter and G. Kotliar, Phys. Rev. Lett. **77**, 131 (1996); A. Martin-Rodero, F. Flores, M. Baldo, and R. Pucci, Solid State Commun. **44**, 911 (1982); A. I. Lichtenstein and M. I. Katsnelson, Cond-mat/9707127 (submitted to Phys. Rev. B).

EFFICIENT AB INITIO TIGHT BINDING

ANDREW HORSFIELD * STEVEN DAVID KENNY **
*Fujitsu European Centre for Information Technology, 2 Longwalk Road, Stockley Park, Uxbridge, UK
**University of Oxford, Department of Materials, Parks Road, Oxford OX1 3PH, UK

ABSTRACT

Tight binding is often seen as a middle ground method, lying between accurate *ab initio* methods, and fast empirical potential methods. The challenge is to make tight binding both as fast and as accurate as possible. One way to achieve this is to take established *ab initio* methods, and apply systematic approximations and efficient numerical techniques to obtain the greatest possible speed. A recently developed method, employing this approach, is described results presented for silicon, and recent developments (including fully self-consistent extensions) are reported.

INTRODUCTION

With the rapid rise in interest in computer modelling at the atomic scale in many disciplines (such as Biology, Chemistry and Materials Science), there is a corresponding need for improved methods. Improvements are required in two directions: accuracy and computational efficiency. Traditionally, if computational efficiency was the top priority, then empirical potentials were used, whereas if accuracy was paramount, *ab initio* quantum mechanical methods were used.

Tight binding offers a successful inbetween way. It is quantum mechanical but also sufficiently simple that it can be applied to complex problems. Of course, tight binding has been used as an electronic structure method for several decades, and has been used to provide qualitative insight into cohesion[1]. Recently, however, its potential as a *quantitative* total energy method has been realised (see, for example, Godwin *et al* [2]).

Because of its place as a middle ground method, it is judged by practitioners of *ab initio* methods in terms of its accuracy, and by practitioners of empirical potential methods in terms of its efficiency. The challenge is to satisfy the requirements of both camps, especially as *ab initio* methods become faster and empirical potentials become more sophisticated.

Empirical tight binding can be as accurate as *ab initio* methods (notably hydrocarbons[3]), and can compete on efficiency grounds with empirical potentials when used with appropriate linear scaling methods. However, its results are sometimes unpredictable, and thus there is a strong incentive to improve on available models.

To create an improved scheme we have chosen to begin from an *ab initio* formalism, as this allows systematic approximations to an accurate theory to be made. The objective is to obtain as much efficiency as possible without compromising unduly on accuracy. An added benefit is the elimination of the fitting of parameters to accurate data. Not only is fitting very time consuming, but it also makes it very difficult to unravel the limitations of the underlying model. The failure of a simulation is always ambiguous: could more accurate results have been obtained by fitting to more data, or is a more sophisticated model called for?

QUICK REVIEW OF AB INITIO TIGHT BINDING METHODS

There have been a number of first principles tight binding methods presented in the literature. Two starting points have been used: linearized muffin tin orbitals (LMTO) and linear combination of atomic orbitals (LCAO).

Andersen and Jepsen[4] presented a way of deriving a two center tight binding Hamiltonian from LMTO. The essential step is to reduce the range of the MTOs, which decay according to a power law, by introducing screening charges which causes the new orbitals to decay exponentially. A linear approximation is then made for the energy dependence of the potential functions, which leads to the tight binding form.

Sankey and Niklewski[5] begin with an LCAO method. The key elements are: the Harris-Foulkes functional[6, 7] is used with the input charge density being a sum of spherical atom-centered charge distributions; the orbitals are compressed and of finite range; the kinetic energy, electrostatic, pseudopotential

and overlap integrals are tabulated; the exchange-correlation integrals are evaluated by means of a nearly-uniform-density approximation. Three center and crystal field terms are included. It is this method which is used as a starting point for the method we present below.

Lin and Harris[8] again use the Harris-Foulkes functional and atomic orbitals. In the original approach analytic functions were used for the orbitals and charge density, so that many integrals could be evaluated analytically, and hence no tables are needed. A quadratic approximation was used for the exchange-correlation integrals. This method has been substantially revised recently in the FastStructure program distributed by MSI.

Stokbro, Chetty, Jacobsen and Norskov[9] developed an approach using effective medium theory with tight binding corrections given by LMTO expressions.

Porezag, Frauenheim, Köhler, Seifert and Kaschner[10] have developed a non-orthogonal tight binding method. While not strictly *ab initio* (it uses non-systematic assumptions) it does use tabulated integrals, and has proven to be very successful for hydrocarbons and semiconductors.

THE NEW SCHEME

The method of Sankey and Niklewski[5] was chosen as the starting point for developing the new method. It was chosen because: the formalism is transparent (all the approximations are clear); it is a flexible approach (there is considerable scope for optimizing basis sets and input charge densities); it is a variational approach (both with respect to input charge density and basis set) thereby minimizing errors; it uses tables for integrals allowing for great computational efficiency; it has been applied successfully to a number of problems.

In order to study sytems containing atoms from the first row of the periodic table some modifications to the original scheme are necessary. The relevant atomic pseudopotentials are deep, so the pseudo charge density has large variations (that is, it is *not* nearly uniform). Thus an alternative scheme to the nearly-uniform-density approximation[5] is needed for the exchange-correlation integrals. Because the pseudopotentials are deep, the pseudo-charge is strongly localised about each nucleus. This suggests using a many-center expansion for the exchange-correlation integrals, which assumes the charge density overlap occurs only between near neighbours. For example, for the exchange-correlation energy the relevant expression would be (including one and two center terms):

$$E_{xc}[\sum_i \rho_i] = \sum_i E_{xc}[\rho_i] + \frac{1}{2} \sum_{i \neq j} \{E_{xc}[\rho_i + \rho_j] - E_{xc}[\rho_i] - E_{xc}[\rho_j]\} + .. \tag{1}$$

where ρ_i is the charge density centered on site i. The first term is a sum of numbers that can be evaluated prior to a simulation. The second term is a sum of pair potentials that can be tabulated prior to a simulation, and then values needed during a simulation can be obtained by interpolation. The same approach applies to the matrix elements of the Hamiltonian, though three center terms must be included in that case.

For molecular dynamics with molecules many different configurations are sampled, so a flexible basis set is called for that can represent the wavefunction accurately for each of these configurations. The virial theorem for particles that interact by means of the Coulomb potential allows us to define some general properties of the basis set. The theorem can be written as $U = -\frac{1}{2}T$, where U is the total energy and T is the kinetic energy, provided there are no net forces present. The more strongly bound the system is, the higher the kinetic energy. The kinetic energy is given by the curvature of the wavefunctions. Thus a flexible basis set is one that can adjust its curvature. A basis set that uses two sets of atomic orbitals (so-called double numeric basis set) turns out to be sufficient. Having two sets of orbitals defines two length scales. A linear combination allows a range of length scales to be sampled. Curvature depends on the inverse of the square of the length scale, thus varying the length scale allows us to vary the curvature. We have tried constructing basis sets in two ways.

In the first way, which has worked very well for molecules, the self-consistent atom problem is solved twice: once for the neutral atom, and once for a positively charged ion (the charge is set at $+2e$ for carbon). The two sets of orbitals generated form the double numeric basis. The single numeric basis is generated from the neutral atom only. From Figure 1 we see that the valence orbitals vary approximately linearly with the net charge on the atom. Thus a linear combination of orbitals taken from the neutral and charged atoms should be able to describe the charge transfer as well as the change in length scale of the orbital with changing environment. The atomic calculations are carried out in a confining potential of

58
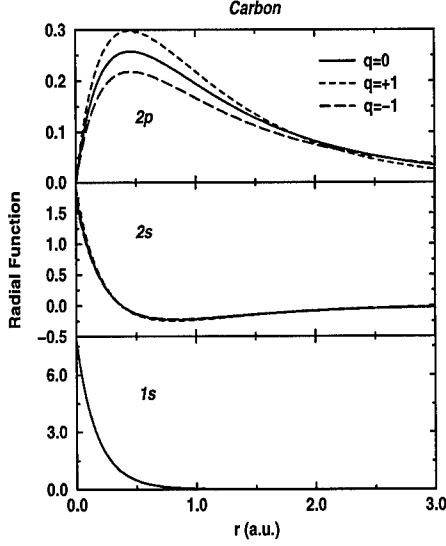
Figure 1: The orbitals for a neutral carbon atom, and for the singly positively and negatively charged ions.

the form $2Ry(r/r_0)^6$ to reduce the range of the orbitals. This is done to reduce the number of neighbors with overlapping orbitals in a simulation. This is very important for calculations involving solids for which there can be very many neighbors.

The second way is the same as the first, except that the second set of orbitals is also obtained from the neutral atom, but using the next higher set of principal quantum numbers, which introduces an extra node into the orbitals[15]. We have not found this to be useful.

As compared with a single numeric basis set, this expanded basis set increases the number of three center integrals by a factor of four. These need to be stored prior to a simulation, and evaluated during a simulation. It is straightforward to condense all the original tables into one set once separable pseudopotentials and the many center expansion for the exchange-correlation terms are used[11], but this leaves the question of how best to carry out the interpolation. Two possible solutions are proposed: one using an expansion in terms of Chebychev polynomials, the other is to use trilinear interpolation. The details of the Chebychev polynomial scheme are given elsewhere[11]. The trilinear interpolation is now explained.

Consider any function of three variables $f(p, q, r)$. Suppose we know the values of this function on a uniform three dimensional grid of points $f_{ijk} = f(p_i, q_j, r_k)$, but wish to know the value of this function at a point (p, q, r) where: $p_i < p < p_{i+1}$, $q_j < q < q_{j+1}$, $r_k < r < r_{k+1}$. The cube that contains the point (p,q,r) thus has corners: (p_i, q_j, r_k), (p_{i+1}, q_j, r_k), (p_i, q_{j+1}, r_k), (p_i, q_j, r_{k+1}), (p_{i+1}, q_{j+1}, r_k), (p_{i+1}, q_j, r_{k+1}), (p_i, q_{j+1}, r_{k+1}), $(p_{i+1}, q_{j+1}, r_{k+1})$. Knowing the values at the corners ($f_{000} = f(p_i, q_j, r_k)$, $f_{100} = f(p_{i+1}, q_j, r_k)$ etc), it is straightforward to carry out a trilinear interpolation to find the value at (p, q, r). Define α_p, β_p, α_q, β_q, α_r and β_r by:

$$
\begin{aligned}
\alpha_p &= \frac{p - p_i}{p_{i+i} - p_i} & \beta_p &= 1 - \alpha_p \\
\alpha_q &= \frac{q - q_j}{q_{j+i} - q_j} & \beta_q &= 1 - \alpha_q \\
\alpha_r &= \frac{r - r_k}{r_{k+i} - r_k} & \beta_r &= 1 - \alpha_r
\end{aligned}
\tag{2}
$$

We then obtain:

$$
\begin{aligned}
f(p, q, r) &\approx \beta_p \beta_q \beta_r f_{000} + \beta_p \beta_q \alpha_r f_{001} + \beta_p \alpha_q \beta_r f_{010} + \alpha_p \beta_q \beta_r f_{100} \\
&+ \beta_p \alpha_q \alpha_r f_{011} + \alpha_p \beta_q \alpha_r f_{101} + \alpha_p \alpha_q \beta_r f_{110} + \alpha_p \alpha_q \alpha_r f_{111}
\end{aligned}
\tag{3}
$$

59

This scheme ensures that the interpolated function is continuous everywhere, though the first derivative in general will not be. However, this expression is extremely fast to evaluate on a computer. Its principle deficiency is that a fine mesh of points is needed for this scheme to give accurate values.

Another bottleneck for the table lookup schemes is the production of the tables in the first place. One approach that we have tried and found to work well is to expand the numerical orbitals as a sum of gaussians. Provided pseudopotentials are used only a few gaussians are needed (typically between 2 and four per orbital). The overlap, kinetic energy and pseudopotential integrals can be evaluated analytically[12]. We use the gaussian pseudopotentials of Goedecker et al [13]. To evaluate the Hartree integrals the density is also expanded in terms of gaussians. The exchange/correlation integrals must be performed numerically.

There are two advantages to this approach: once the gaussians have been fitted there is no further error in the integral tables (at least in the absence of exchange and correlation integrals), thus all the integrals are internally consistent with one another; the analytical integrals can be evaluated in minutes rather than hours, making building tables much more convenient (the exchange and correlation integrals are still slow, however).

There is one important point that needs to be taken care of, however. Gaussians do not go to zero at some finite cutoff, so care has to be taken to ensure that the gaussian fits have a very small value at the cutoff radius used to generate the numerical orbitals. One way to achieve this is to use a harmonic potential $(V(r) = V_0 r^2)$ to confine the the atom in, and then adjust the prefactor (V_0) to make the orbitals decay at the appropriate rate. This form is chosen because the wavefunctions for a harmonic oscillator are gaussian in form, so the fit becomes more natural, especially in the tail region. This is found to remove most of the tendency for the fitted orbitals to project beyond the numerical ones.

The following analysis allows us to estimate what value V_0 should take. Note that all lengths are in Bohr radii, and all energies are in Rydbergs. The ground state of the harmonic oscillator has the wavefunction $\exp(-\alpha r^2)$, where $\alpha = \frac{1}{2}\sqrt{V_0}$. If r_{cut} is the cutoff radius, and we wish $\exp(-\alpha r_{cut}^2) < \varepsilon$, then $\alpha > \ln(1/\varepsilon)/r_{cut}^2$. Therefore

$$V_0 > \left[\frac{2\ln(1/\varepsilon)}{r_{cut}^2}\right]^2 . \tag{4}$$

In practice we have found that $\varepsilon = 10^{-3}$ works well. This may seem a large value, but the atomic core tends to pull the wavefunction in, so the actual wavefunctions tend to be well below ε at r_{cut}.

The effect on the geometries and energies of hydrocarbons of introducing the second set of orbitals into the basis set has already been demonstrated[11]. It was found that a double numeric basis set gave significantly improved geometries compared with single numeric basis sets, whereas the influence of polarization orbitals (typically d-orbitals) is small.

The effect on the band structure of silicon of introducing the second set of basis functions can be seen in Figure 2. The valence band is essentially unaffected by the second set of orbitals. However, the location and size of the band gap is greatly improved. Thus we see that a double numeric basis set is important for a reliable description of the electronic structure and thus for the description of defects etc. However, this basis set gives a lattice constant that is too large (about 5.55Å, compared with the experimental value of 5.43Å, and the planewave value of 5.39Å[16]), and a bulk modulus that is too small (0.83Mbars, compared with the experimental value of 0.99Mbars and the planewave value of 0.94Mbars[16]). We believe that d-orbitals are necessary to obtain improved results.

SELF-CONSISTENT EXTENSIONS

It is not without irony that the project of turning fully self-consistent ab initio methods into approximate methods has moved on and become a project to produce very efficient self-consistent methods. This development has been pioneered by Ordejon, Artacho and Soler[14].

The central idea is to take the Harris-Foulkes Hamiltonian from the tables and then add in corrections due to the charge redistribution that occurs when self-consistency is imposed. Because no shape is assumed for the self-consistent charge density it must be evaluated on a numerical grid. A uniform grid is used as this allows Fast Fourier Transforms to be used to evaluate the correction to the Hartree potential. Because the changes in the charge density are quite small, a coarse mesh can be used. This coupled with the fact that we only need to transform this once (rather than for each k-point) makes the method more efficient than plane wave methods. A second important ingredient is that the energy and forces can be found using a linear scaling algorithm - which is impossible with current plane wave methods.

Single Numeric   Double Numeric
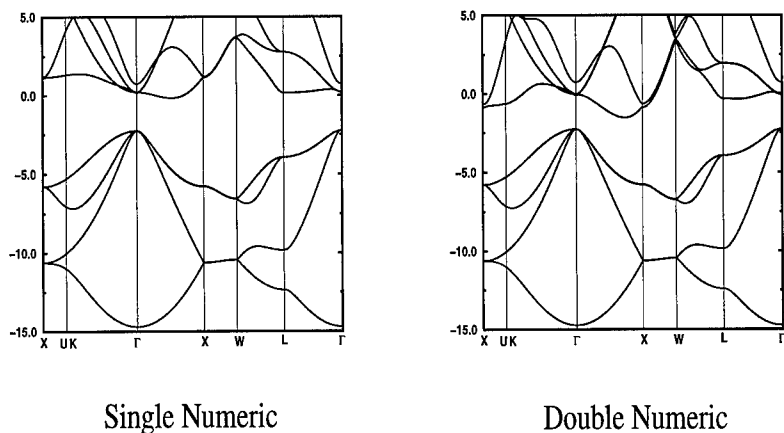
Figure 2: The band structure for silicon evaluated using single and double numeric basis sets. Note that the single numeric basis set give the correct valence band, but introduces errors into the conduction band.
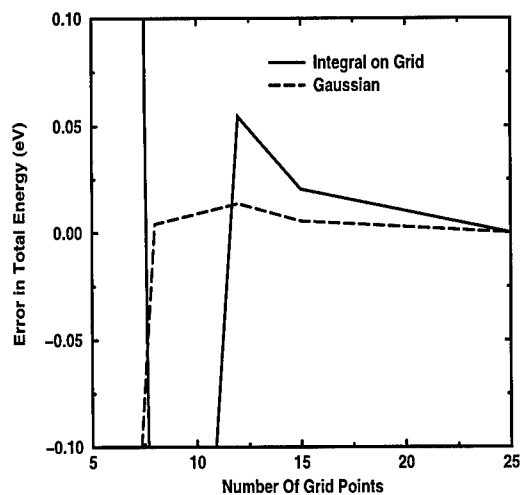
Figure 3: The error in the total energy of bulk silicon evaluated using the two methods for avoiding three center integral tables. Note that the x axis is the number of grid points on one side of the cell. Thus the total number of points on the grid is the cube of this.

61
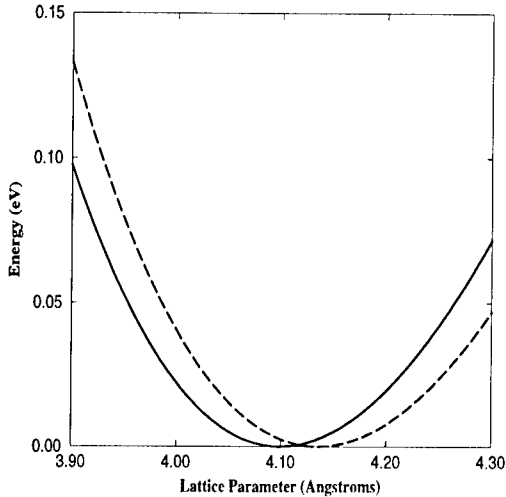
Figure 4: The bulk modulus curve for aluminum, the solid line is the result of a calculation using the method described above the dashed line a plane wave pseudopotential calculation.

Three center integrals tables can be large and slow to construct. Thus it is desirable to eliminate them. Two methods have been proposed to achieve this. To understand them we need to note that for self-consistent calculations the tables contain integrals of the form:

$$I_{i\alpha,j\beta,k} = \int \phi_{i\alpha}(\vec{r} - \vec{R}_i) V_k^{NA}(\vec{r} - \vec{R}_k) \phi_{j\beta}(\vec{r} - \vec{R}_j) d\vec{r} \tag{5}$$

where $V_k^{NA}(\vec{r} - \vec{R}_k)$ is the so-called neutral atom potential. It consists of the sum of the local part of the pseudopotential and the Hartree potential of the spherical charge density at that site. The approach of Sanchez-Portal et al[15] is to perform these integrals on a uniform mesh along with the exchange-correlation and correction to the Hartree potential. The one we have developed uses gaussian expansions for the orbitals and the potential. We can then use analytic results for the integrals that can be evaluated very efficiently. Results obtained for the total energy of bulk silicon as a function of the number of mesh points using both methods are shown in Figure 3. The errors in both cases are small. However, the convergence of the numerical integral method is sufficiently slow to make the evaluation of quantities that depend on small energy differences (such as elastic constants) very expensive.

We have used this method to perform some calculations for aluminum. The results for the bulk modulus curve are extremely good. The predicted lattice parameter is 4.10Å and the bulk modulus 0.72Mbar which is in excellent agreement with the experimental values of 4.05Å and 0.76Mbar. The calculated bulk modulus curve is given in Figure 4. The solid line is the result found using the method detailed above and the dashed line is the result of plane wave pseudopotential calculations (for which the lattice parameter is 4.14Å and the bulk modulus is 0.68Mbar). It should be noted that the plane wave calculation used 50 times as much memory, and took 10 times as long to complete, compared with the localised orbital calculation.

CONCLUSION

A recently developed *ab initio* tight binding scheme has been described in outline. The benefits of using a double numeric basis set are described, and seen to span total energies, geometries, and electronic structure. A highly efficient trilinear interpolation scheme has been described, as have some recent

62

developments to make the method fully self-consistent.

ACKNOWLEDGEMENTS

REFERENCES

1. D. Pettifor, *Bonding and Structure of Molecules and Solids*, (Clarendon Press, Oxford, 1985)

2. P. D. Godwin, A. P. Horsfield, D. G. Pettifor, and A. P. Sutton, Phys. Rev. B **54**, 15776 (1996)

3. A .P. Horsfield, P. D. Godwin, D. G. Pettifor and A. P. Sutton, Phys. Rev. B **54**, 15773 (1996)

4. O. K. Andersen and O. Jepsen, Phys. Rev. Lett. **53**, 2571 (1984)

5. O. F. Sankey and D. J. Niklewski, Phys. Rev. B **40**, 3979 (1989)

6. J. Harris, Phys. Rev. B **31**, 1770 (1985)

7. W. M. C. Foulkes, Ph.D. Thesis, University of Cambridge (1987)

8. Z. Lin and J. Harris, J.Phys.: Condense. Matter **4**, 1055 (1992)

9. K. Stokbro, N. Chetty, K. W. Jacobsen and J. K. Norskov, Phys. Rev. B **50**, 10727 (1994)

10. D. Porezag, Th. Frauenheim, Th. Kohler, G. Seifert and R. Kaschner, Phys. Rev. B **51**, 12947 (1995)

11. A. P. Horsfield, Phys. Rev. B **56**, 6594 (1997)

12. S. F. Boys, Proc. Royal. Soc. **A200**, 542 (1950)

13. S. Goedecker, M. Teter, J. Hutter, Phys. Rev. B **54**, 1703 (1996)

14. P. Ordejon, E. Artacho and J. M. Soler, Phys. Rev. B **53**, 10441 (1996)

15. D. Sanchez-Portal, P. Ordejon, E. Artacho and J. M. Soler, *To be published*

16. D. Allan, Private Communication (1997)

EFFECTIVE INTERATOMIC INTERACTIONS VIA THE TB-LMTO METHOD

V. DRCHAL[a,b], J. KUDRNOVSKÝ[a,b], A. PASTUREL[c], I. TUREK[d], P. WEINBERGER[b],
A. GONIS[e], and P.E.A. TURCHI[e]

[a]Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2,
 CZ-180 40 Praha 8, Czech Republic
[b]Institute for Technical Electrochemistry, Technical University of Vienna, Getreidemarkt 9,
 A-1060 Vienna, Austria
[c]Experimentation Numérique, Maison des Magistères, CNRS, BP 166, 38042 Grenoble
 Cedex, France
[d]Institute of Physics of Materials, Academy of Sciences of the Czech Republic, Žižkova 22,
 CZ-616 62 Brno, Czech Republic
[e]Lawrence Livermore National Laboratory, Livermore, CA 94550, U.S.A.

ABSTRACT

The energetics of metallic alloys, their surfaces or interfaces, and magnetic multilayers
is studied in terms of effective interatomic (or interlayer) interactions that are determined
from ab initio electronic structure calculations using the TB-LMTO method combined with
the coherent potential approximation and the method of surface Green functions. First
the theoretical background (force theorem, Lloyd formula, generalized perturbation method
for bulk and surfaces, vertex cancellation theorem, method of infinitesimal rotations) is
discussed, and then the applications to the phase stability of bulk alloys, surface segregation
in disordered alloys, magnetism-induced ordering in two- and three-dimensional systems,
phase diagram of two-dimensional alloys, interlayer exchange coupling in metallic multilayers,
and the construction of Heisenberg-like Hamiltonians for magnetic systems are presented.

INTRODUCTION

The thermodynamical properties of various systems such as alloys, their surfaces, mag-
netic multilayers, and ferromagnets can be predicted on an ab initio level. First, the internal
energy and its dependence on the configuration of the system is expressed in terms of an
effective Hamiltonian. The parameters of this Hamiltonian (which are often called effective
cluster interactions (ECIs)) are determined from first principles electronic structure calcula-
tions. In the second step, the thermodynamical properties of the system are studied by the
methods of statistical mechanics.

By using an effective Ising-type Hamiltonian the configurational dependence of the total
energy of a disordered binary alloy $A_x B_{1-x}$ can be expressed as

$$H_I = E_0 + \sum_{\mathbf{R}} D_{\mathbf{R}} \, \eta_{\mathbf{R}} + \frac{1}{2} \sum_{\mathbf{RR'}} V_{\mathbf{RR'}} \, \eta_{\mathbf{R}} \, \eta_{\mathbf{R'}} + \dots . \tag{1}$$

The parameters entering (1) are the configurationally independent part of the total energy
E_0, the on-site energies $D_{\mathbf{R}}$, the interatomic pair nteractions $V_{\mathbf{RR'}}$, and generally, the inter-
atomic interactions of higher order. A particular configuration of the alloy is determined by
a set of occupation indices $\eta_{\mathbf{R}}$, where $\eta_{\mathbf{R}} = 1$ if the site $\mathbf{R}$ is occupied by an atom of the
type A, and $\eta_{\mathbf{R}} = 0$ otherwise.

By using an effective (classical) Heisenberg-type Hamiltonian the total energy of a ferro-
magnet can be expressed as

$$H_H = E_0 - \sum_{\mathbf{RR'}} J_{\mathbf{RR'}} \, \mathbf{e_R} \cdot \mathbf{e_{R'}} + \dots , \tag{2}$$

65

where $\mathbf{e_R} = \mathbf{m_R}/|\mathbf{m_R}|$ is the unit vector in the direction of a magnetic dipole $\mathbf{m_R}$ at site $\mathbf{R}$.

THEORY

The tight-binding linear muffin-tin orbital method

Our approach [1, 2, 3] is based on the local spin density approximation to the density functional formalism. The electronic structure is calculated using the first-principles all-electron linear muffin-tin orbital (LMTO) method [4, 5] in its tight-binding version [6, 7] (TB-LMTO) and assuming the atomic sphere approximation (ASA).

Substitutional chemical disorder is included within the coherent potential approximation (CPA) [8] which describes reliably concentration dependent trends of the electronic structure in bulk alloys and at their surfaces or interfaces.

The effect of layered geometry of surfaces or multilayers is treated via the surface Green function formalism [9]. The charge and spin density self-consistency is achieved within the ASA, but non-spherical terms of the charge densities are used in the construction of the interatomic Coulomb interactions in the surface or interface layers. It was shown [10, 11, 12] that a correct description requires the inclusion of net charges and dipole moments in the surface/interface region. The electrostatic potential barrier across the solid-vacuum or solid-solid interface (and the closely related work function) as well as the surface/interface energy sensitively depend on the details of the charge distribution as given by the layer dependence of the atomic net charges and the non-spherical character of the charge distribution inside the atomic spheres. The method substitutes to some extent the full-potential techniques and reliably describes surfaces and interfaces with charge densities that strongly deviate from spherical symmetry. An important advantage of this approach is the treatment of the surface and bulk atomic layers on an equal footing which is important for an accurate determination of small energy differences needed in the construction of effective Hamiltonians.

The method is well suited to study spin-polarized systems and to include relativistic effects relevant for the energetics of systems containing heavy noble and transition metals [1, 13, 14].

Effective Hamiltonians

The parameters of the effective Hamiltonians are determined by mapping the total energy of various configurations of the system onto the effective Hamiltonian. A direct approach is represented by the Connolly-Williams inversion scheme [15] that extracts the interatomic interactions in alloys from total energies of several ordered structures. A similar approach can be used also for construction of an effective Heisenberg-type Hamiltonian. Here we follow a different approach based on the knowledge of the ab initio electronic structure of a reference configuration Γ_0 of the system. The energies of other configurations are calculated perturbatively using the force theorem [16] without repeating the self-consistent calculation. The total energy $E_{tot}[\Gamma, \rho(\Gamma)]$ of a system in the ground state that corresponds to the configuration Γ is a functional of the charge (and spin) density $\rho(\Gamma)$ and attains its minimum in the ground state. Consequently, it is stationary with respect to small changes in $\rho(\Gamma)$. The charge (and spin) density $\rho(\Gamma_0)$ corresponding to the reference configuration Γ_0 can be used to construct the approximate charge (and spin) density $\tilde{\rho}(\Gamma)$ for the other configuration Γ. The approximate $\tilde{\rho}(\Gamma)$ differs from the exact $\rho(\Gamma)$ by $\delta\rho$. Due to the stationary property of $E_{tot}[\Gamma, \rho(\Gamma)]$, the error in energy is of order $||\delta\rho||^2$.

In the alloy case the approximate $\tilde{\rho}(\Gamma)$ is constructed by using the charge densities $\rho_{\mathbf{R}}^Q$ corresponding to the atomic species Q occupying site $\mathbf{R}$. The quantities $\rho_{\mathbf{R}}^Q$ are in turn used to construct the potentials in each atomic sphere, then the potential parameters, and finally the one-electron Hamiltonian. The charge densities, atomic potentials and potential para-

meters are identical for all equivalent atoms, while the fluctuations induced by individually different environments of atoms are neglected.

In the case of magnetic systems the configuration Γ is usually generated from the reference configuration Γ_0 by changing one, or several local spin quantization axes. The new $\bar{\rho}(\Gamma)$ is then constructed by applying corresponding operations of rotation to $\rho_{\mathbf{R}\sigma}$.

The change of configuration always leads (via changes in the one-electron LMTO Hamiltonian) to changes in the band term of the total energy, and, depending on the system studied, it could require also corrections to core, double-counting, electrostatic, and magnetostatic terms.

A suitable tool to calculate the change in the band term is the Lloyd formula [17]

$$\Delta\Omega_{band} = \Omega_{band}[\Gamma] - \Omega_{band}[\Gamma_0] = \frac{1}{\pi}\operatorname{Im}\int_{-\infty}^{\infty} f(z)\operatorname{Tr}\ln\left[1 - VG^{(0)}(z)\right]dE\,, \tag{3}$$

where $z = E + i0$, $f(z)$ is the Fermi-Dirac function, $G^{(0)}(z)$ is the Green function of the reference configuration, and V is the operator of perturbation that transforms configuration Γ_0 into Γ. As usual, a grand canonical ensemble is assumed in (3) which leads to a replacement of the energy by the grand canonical potential. The main advantage of the Lloyd formula (3) comes from the partitioning of $\operatorname{Tr}\ln$. Let P and Q be two complementary projection operators, $\mathsf{P} + \mathsf{Q} = 1$, then

$$\operatorname{Tr}\ln A = \operatorname{Tr}_{\mathsf{P}}\ln\left[A_{\mathsf{PP}} + A_{\mathsf{PQ}}\frac{\mathsf{Q}}{A_{\mathsf{QQ}}}A_{\mathsf{QP}}\right] + \operatorname{Tr}_{\mathsf{Q}}\ln A_{\mathsf{QQ}}\,. \tag{4}$$

If the perturbation V acts on the subspace P, the equation (3) simplifies to the form

$$\Delta\Omega_{band} = \frac{1}{\pi}\operatorname{Im}\int_{-\infty}^{\infty} f(z)\operatorname{Tr}_{\mathsf{P}}\ln\left[\mathsf{P} - V_{\mathsf{PP}}G_{\mathsf{PP}}^{(0)}(z)\right]dE\,, \tag{5}$$

i.e., it can be evaluated only in the subspace P. This formulation is much more efficient than that based on the local densities of states (LDOS) as the changes in LDOSs induced by the local perturbation extend over the whole space. For example, for the calculation of the pair interaction between two atoms in the alloy, the subspace P is spanned just by the orbitals on these two sites.

The band term is represented by the Lloyd formula and the mapping onto an effective Hamiltonian is given by expanding the Lloyd formula in terms of relevant variables (occupation indices $\eta_{\mathbf{R}}$ in the alloy case, spin variables for magnetic systems).

APPLICATIONS

Alloy phase stability

The parameters of the alloy Ising-type Hamiltonian (1) are determined by the generalized perturbation method (GPM) [18, 19, 20] (for its implementation within the LMTO method see [1, 21, 22]). The GPM is based on the Lloyd formula (3) (at $T = 0$ K) and it is an expansion around the CPA effective medium that represents the reference state. The importance of the CPA for the GPM is due to the variational properties of the band energy term [23]. The integrated density of states, and consequently, the grand canonical potential and also the band energy term evaluated within the CPA are stationary with respect to variations of the effective medium. The condition of a vanishing first variation is equivalent to the CPA equation for the effective medium.

The parameters $D_{\mathbf{R}}$ and $V_{\mathbf{RR'}}$ of the alloy Hamiltonian (1) are given as

$$D_{\mathbf{R}}^{band} = D_{\mathbf{R}}^{A,band} - D_{\mathbf{R}}^{B,band}\,, \tag{6}$$

where

$$D_{\mathbf{R}}^{Q,band} = -\frac{1}{\pi} \operatorname{Im} \int_{-\infty}^{E_F} \operatorname{tr}_{L\sigma} \ln \left[1 + \left(P_{\mathbf{R}}^{Q}(z) - \mathcal{P}_{\mathbf{R}}(z) \right) \bar{g}_{\mathbf{RR}}(z) \right] dE. \tag{7}$$

Here, $P_{\mathbf{R}}^{Q}(z)$ is the potential function, $\mathcal{P}_{\mathbf{R}}(z)$ is the coherent potential function, $\bar{g}_{\mathbf{RR}}(z)$ is the on-site element of the auxiliary Green function, and $\operatorname{tr}_{L\sigma}$ denotes the trace over the orbital momentum $L = (\ell m)$ and spin σ. The (unrenormalized) pair interactions are then

$$V_{\mathbf{RR}'}^{band} = \frac{1}{\pi} \operatorname{Im} \int_{-\infty}^{E_F} \operatorname{tr}_{L\sigma} \tau_{\mathbf{R}}(z) \bar{g}_{\mathbf{RR}'}(z) \tau_{\mathbf{R}'}(z) \bar{g}_{\mathbf{R}'\mathbf{R}}(z) dE, \tag{8}$$

where $\tau_{\mathbf{R}}$ is the difference of t-matrices for A and B atoms. For case studies see [1, 24]. We note that this formalism can be generalized also to multicomponent alloys [25]. The contributions from other terms of the total energy (double counting, core and electrostatic terms) are important at alloy surfaces [26].

Once the ECIs are known, the methods of statistical mechanics can be used to determine the thermodynamical potential of the alloy. In this way, it is possible to search for the most stable structure of the system at a given temperature T and thus to derive its phase diagram from the first principles. Various methods are used. The most simple is the single-site effective medium approximation (Bragg-Williams approximation) which might be satisfactory at high temperatures, but which often overestimates the transition temperatures. A substantial improvement was achieved by introducing the cluster variation method (CVM) that properly includes the multisite correlations [19]. Highly reliable results are obtained by Monte Carlo simulations (see e.g. [27]). The Monte Carlo method was applied to bulk alloys, alloy surfaces [22, 26], and to studies of magnetism-induced ordering [28]. In the case of concentrated alloys, the pair interactions $V_{\mathbf{RR}'}$ decay exponentially with increasing distance $|\mathbf{R} - \mathbf{R}'|$ due to the imaginary part of the CPA self-energy entering $\bar{g}_{\mathbf{RR}'}(z)$. This fact facilitates the application of the Monte Carlo technique as its numerical efficiency strongly decreases with the number of nonzero pair interactions.

Surface segregation

The segregation of one species at alloy surfaces has been the subject of extensive experimental and theoretical studies (see [29] for a review). It plays an important role in such diverse phenomena as catalysis, chemisorption, and crystal growth since all of them depend sensitively on surface properties.

The surface energetics can be treated on two levels of sophistication:
(i) Truncated bulk approximation [21, 22, 30, 31]. The single-site charge density is that of the homogeneous bulk alloy and the redistribution of the electronic charge in the surface region is not considered. The atomic potentials are identical to those of the homogeneous bulk alloy, and the CPA effective medium of the homogeneous bulk alloy is used as a reference medium in the GPM. The ECIs are determined by mapping only the band term onto the Ising Hamiltonian. The core, double-counting, Madelung and non-spherical terms do not contribute.
(ii) Selfconsistent theory [26, 32, 33]. The single-site charge density, atomic potentials, and the CPA effective medium are determined in a fully self-consistent calculation of the surface electronic structure including the charge redistribution in the surface region. Besides the band term, also the core, double-counting, Madelung and non-spherical terms contribute to the ECIs. In contrast to the previous simplified formulation, the Ising Hamiltonian parameters can be determined for any concentration profile of the surface region.

This allows to determine in a fully self-consistent manner the surface concentration profile. Two steps, namely, the calculation of the ECIs for a given concentration profile, and the

determination of the new concentration profile using the methods of statistical mechanics for a given set of ECIs, are repeated until a full consistency between the electronic structure and concentration profile is achieved.

The equilibrium composition of the surface layers is simulated using the Monte Carlo technique [22, 26].

Using the fully self-consistent procedure, we have studied the surface segregation in Cu-Ni, Ag-Pd, and Pt-Rh alloy systems. As an illustration we show here the results for the (001) face of the fcc $Cu_{25}Ni_{75}$ (Fig. 1) and $Pt_{55}Rh_{45}$ (Table I) alloys. The calculated concentration profiles for Cu-Ni alloys are in a good agreement with accessible experimental data which show a quite pronounced tendency to segregation of Cu and damped concentration oscillations in subsurface layers [26].
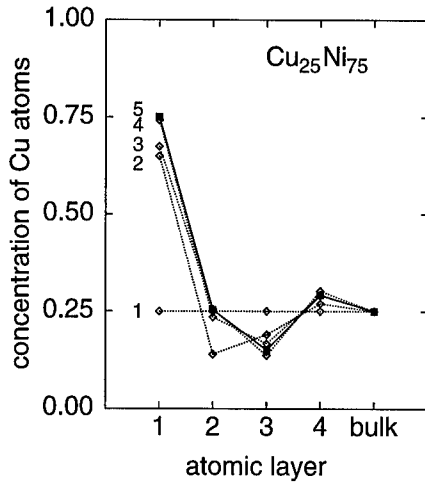
Figure 1: Evolution of the segregation profile in the first 5 iteration steps for the (001) face of the fcc $Cu_{25}Ni_{75}$ alloy.

Table I: Selfconsistent determination of the segregation profile at the (001) surface of the fcc $Pt_{55}Rh_{45}$ alloy. Concentrations $c_p(Pt)$ of Pt atoms in the first 4 layers as obtained in the first 4 iteration steps are compared to experimental data.

iteration	$c_1(Pt)$	$c_2(Pt)$	$c_3(Pt)$	$c_4(Pt)$
1	0.80	0.35	0.65	0.45
2	0.76	0.35	0.62	0.48
3	0.78	0.39	0.64	0.48
4	0.77	0.37	0.64	0.47
experiment[a]	0.78	0.42	0.64	0.47

[a]Ref. [34]

The surface segregation for Pt-Rh was studied both experimentally [34] and theoretically [35]. It is now agreed that Pt segregates to the surface upon annealing at temperatures above 900 K and the concentration profile has oscillatory character. The Monte Carlo simulations were done at 973 K, i.e., at the same temperature at which the experimental study was

performed [34]. The self-consistent calculations of the segregation profile were started from an oscillatory profile close to the experimental one. Usually 3 or 4 iterations were needed to achieve self-consistency with an accuracy of order 1 % between the previous and new concentration profile. Also the changes in the Ising Hamiltonian parameters between two consecutive iterations rapidly converged to zero. The agreement between the calculated and experimental profile is remarkably good (see Table I), except for the second layer, where the calculated value (0.37) differs from the experimental one (0.42), but still lies inside the experimental error bar.

Two-dimensional alloys

The phase diagram of a two-dimensional alloy system formed by a random overlayer on a perfect nonrandom substrate can be studied using similar methods as for bulk alloys or surfaces. Such system is truly two-dimensional from the point of view of statistical mechanics, while its electronic structure is that of a semi-infinite three-dimensional system.

Table II: Pair interactions $V_n(x)$ (in mRy) up to 6-th neighbors for a two-dimensional Cu_xPd_{1-x} random alloy on a perfect fcc(001) Cu substrate for 3 concentrations $x = 0.25$, 0.5, and 0.75, and for the hypothetical random alloy $Pd_{50}(vac)_{50}$. All the other pair and triplet interactions are less than 0.15 mRy.

n	1	2	3	4	5	6
$V_n(Cu_{25}Pd_{75})$	5.59	−1.10	−0.22	−0.20	−0.02	0.22
$V_n(Cu_{50}Pd_{50})$	6.51	−0.59	−0.36	−0.18	−0.13	−0.01
$V_n(Cu_{75}Pd_{25})$	7.70	0.04	−0.07	0.12	0.00	−0.06
$V_n(Pd_{50}(vac)_{50})$	−19.79	1.38	−0.39	−0.03	−0.04	0.23
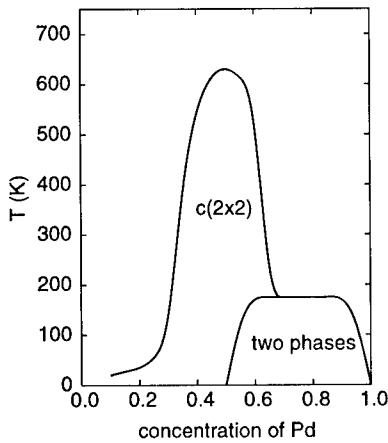
Figure 2: Phase diagram of the Cu-Pd surface alloy on a Cu (001) substrate.

As an example we show here the results for a Cu-Pd random overlayer on the fcc(001) Cu substrate. The pair interactions on a two-dimensional square lattice are reported in Table II. The calculations for a hypothetical random two-dimensional alloy $Pd_{50}(vac)_{50}$ rule out

70

the possibility that the overlayer is formed by some ordered structure of Pd atoms only. Large negative values of pair interactions between the Pd atoms and empty spheres denoted here as (vac) indicate a strong tendency for clustering of Pd atoms on the Cu surface. On the other hand, the positive pair interactions found for Cu_xPd_{1-x} random overlayer suggest an ordering tendency [36]. This was confirmed by extensive Monte Carlo simulations for chemical composition and temperature varying in a broad range [37] (Fig. 2). The ordered phase $c(2 \times 2)$ is stable for $0.35 < x < 0.65$ at elevated temperatures, the critical temperature of the second-order phase transition at $x = 0.5$ is above 600 K. The stability range of the ordered phase is relatively large at room temperature. On the Pd-rich side a first-order transition occurs below 200 K corresponding to a mixture of the $c(2 \times 2)$ ordered phase and pure Pd. It should be emphasized that the phase diagram is asymmetric with respect to $x = 0.5$, which is a consequence of the concentration dependence of pair interactions. These theoretical results are in a good agreement with experimental data.

Magnetism-induced ordering

The magnetic order can substantially influence atomic ordering in alloys. We show that the ordering in the bulk bcc $Fe_{50}Co_{50}$ alloy arises from the presence of ferromagnetism in this system. According to the experimental data [38, 39], the Curie temperature of the $Fe_{50}Co_{50}$ alloy $T_C = 1255$ K lies above the ordering temperature $T_{ord} = 1003$ K. Below T_{ord} the alloy has the ordered CsCl (B2) structure, while for intermediate temperatures, $T_{ord} < T < T_C$ it forms a bcc solid solution, and at T_C it undergoes a martensitic transformation into the (non-magnetic) fcc structure which is stable up to the melting point ($T_C < T < T_m$). In order to elucidate the mutual interplay of magnetism and ordering in this alloy, we have calculated [28] the pair interactions V_n up to 11th neighbors both for the paramagnetic and the ferromagnetic state of the alloy (Table III). The pair interactions in the paramagnetic state would give a weak tendency to phase separation. In the magnetic state, the majority band ($\sigma = \uparrow$) is almost fully occupied and its contribution is negligibly small (closed shells do not contribute to chemical bonding). The minority band ($\sigma = \downarrow$) moves to higher energy and contributes substantially to the pair interactions which now favor attraction of different atomic species in the first coordination sphere leading thus to ordering.

Table III: Pair interactions (in mRy) up to 6-th neighbors for the bcc based $Fe_{50}Co_{50}$ alloy in the paramagnetic (V_n^{para}) and in the ferromagnetic state ($V_n^{ferro} = V_n^{\uparrow} + V_n^{\downarrow}$).

n	1	2	3	4	5	6
V_n^{para}	−0.68	−0.07	−0.16	−0.08	0.14	0.06
$V_n^{\uparrow}$	0.01	0.03	0.01	0.00	0.00	0.00
$V_n^{\downarrow}$	1.96	−0.66	0.21	0.08	−0.31	0.05
V_n^{ferro}	1.97	−0.63	0.22	0.09	−0.31	0.05

The stability of a random phase with respect to the formation of an ordered structure is, within a simple mean-field theory, related to the Fourier transform $V(\mathbf{k})$ of the pair interactions $V_{\mathbf{RR'}}$ [40]. An absolute minimum of $V(\mathbf{k})$ at $\mathbf{k}_0 = 0$ indicates the tendency to phase separation, while absolute minima at $\mathbf{k}_0 \neq 0$ indicate an ordering tendency to form a superstructure compatible with the vector $\mathbf{k}_0$. A mean-field estimate for the ordering temperature depends on $V(\mathbf{k}_0)$

$$T_{ord} = -x(1 - x)V(\mathbf{k}_0)/k_B \,, \tag{9}$$

where k_B is the Boltzmann constant. Including 11 pair interactions, we find $T_{ord} = 917$ K for Fe$_{50}$Co$_{50}$ alloy. The Monte Carlo simulation yields a T_{ord} lower approximately by 20 %, but from technical reasons, it is limited to a smaller number of pair interactions.

A similar analysis can be done also for Fe$_{50}$Co$_{50}$ overlayer on a Cu (001) substrate [28]. The pair interactions calculated in the paramagnetic state show a tendency to phase separation, while those computed in the ferromagnetic state correspond to ordering into a $c(2 \times 2)$ structure. In comparison with Monte Carlo results, the mean-field theory overestimates the ordering temperature T_{ord} by about 40 %, which is not surprising in view of the well-known fact that the accuracy of a mean-field theory for two-dimensional systems is rather limited. For example, if the first 5 pair interactions are included, the mean-field estimate is $T_{ord} \approx 560$ K, while the Monte Carlo method yields $T_{ord} \approx 380$ K. Unfortunately, experimental data are not available for comparison.

In this section, we have presented a very simple theory of the mutual interplay between the chemical order and magnetism in magnetic alloys. These two phenomena should be treated on the same footing as proposed earlier by Bieber and Gautier [41].

Interlayer exchange coupling

Interlayer exchange coupling occurs in many multilayer systems composed of magnetic and non-magnetic layers, and it is in some cases accompanied by an oscillatory magnetoresistance.

Recently, ab initio formulations of the interlayer exchange coupling have appeared [42, 43, 44], based on an application of a layer version of the Lloyd formula in order to evaluate the difference between the grand canonical potentials of the ferromagnetic (F, $\theta = 0$) and antiferromagnetic (AF, $\theta = \pi$) alignment using the so-called frozen potential approximation [42].

The system considered consists of a central finite non-magnetic spacer slab (C) of varying thickness N and two semi-infinite systems, denoted $\mathcal{L}$ (left) and $\mathcal{R}$ (right), each containing M magnetic layers on top of a semi-infinite non-magnetic spacer. We assume that the spin orientation in the right magnetic slab $\mathcal{R}$ is rotated by an angle θ with respect to that of the left magnetic slab $\mathcal{L}$.

The quantity of physical interest is the difference of the grand canonical potentials between the ferromagnetic ($\theta = 0$) and a rotated ($\theta \neq 0$) alignment of the two magnetic slabs, namely the exchange energy $\mathcal{E}_x(\theta) = \Omega(\theta) - \Omega(0)$. The exchange energies $\mathcal{E}_x(\theta)$ are very small quantities and their direct evaluation is rather cumbersome [45]. The clue is to use the Lloyd formula in the framework of the frozen potential approximation as applied to a rotated alignment which differs from the ferromagnetic alignment by a localized perturbation [42, 43, 44, 46]. Technically there are a few possibilities for how to partition the system into an unperturbed part and a localized perturbation. A common approach [42, 44, 46] is to consider the rotated magnetic slab as a perturbation. Clearly, with increasing thickness M of the magnetic slabs the numerical effort increases as the third power of M. This limitation can be relaxed [43] by considering three decoupled non-interacting regions $\mathcal{L}$, C , and $\mathcal{R}$ as an unperturbed system. The localized perturbation is the interlayer coupling at the $\mathcal{L}/C$ and the $C/\mathcal{R}$ interfaces, which is independent of the thickness of the magnetic slabs. The concept of principal layers (PL) [9] as used within the TB-LMTO method leads to a block tridiagonal form of the structure constants and of the inverse of the Green function. Employing partitioning technique to the trace of the logarithm of the Green function (4), it is possible to extract directly the term describing the coupling of interfaces, $\delta \operatorname{Tr} \ln \mathcal{G}(z)$,

$$\delta \operatorname{Tr} \ln \mathcal{G}(z) = -\frac{1}{N_{\|}} \sum_{\mathbf{k}_{\|}} \operatorname{tr}_{\mathrm{L}\sigma} \ln \left[1 - \tau_1(\mathbf{k}_{\|}, z) \, g_{1N}(\mathbf{k}_{\|}, z) \, \tau_N(\mathbf{k}_{\|}, z) \, g_{N1}(\mathbf{k}_{\|}, z) \right] , \qquad (10)$$

where the auxiliary Green function $\mathcal{G}(z) = [P(z) - S]^{-1}$ is defined in terms of the potential function $P(z)$ and structure constant matrix S, and the t-matrices τ_i $(i = 1, N)$

$$\tau_i(\mathbf{k}_{\parallel}, z) = \Gamma_i(\mathbf{k}_{\parallel}, z) \left[1 - g_{ii}(\mathbf{k}_{\parallel}, z) \, \Gamma_i(\mathbf{k}_{\parallel}, z)\right]^{-1} \tag{11}$$

are defined in terms of the embedding potentials at the interfaces,

$$\Gamma_1(\mathbf{k}_{\parallel}, z) = S_{10}(\mathbf{k}_{\parallel}) \, \tilde{\mathcal{G}}(\mathbf{k}_{\parallel}, z) \, S_{01}(\mathbf{k}_{\parallel}), \quad \Gamma_N(\mathbf{k}_{\parallel}, z) = S_{01}(\mathbf{k}_{\parallel}) \, \mathcal{G}(\mathbf{k}_{\parallel}, z) \, S_{10}(\mathbf{k}_{\parallel}). \tag{12}$$

The quantities $g_{11}(z), g_{1N}(z), g_{N1}(z)$, and $g_{NN}(z)$ are blocks of the spacer slab Green function consisting of N layers, $S_{01}(\mathbf{k}_{\parallel})$ and $S_{10}(\mathbf{k}_{\parallel})$ are the interlayer structure constants which couple neighboring PLs, and $\tilde{\mathcal{G}}(\mathbf{k}_{\parallel}, z)$ and $\mathcal{G}(\mathbf{k}_{\parallel}, z)$ are the surface Green functions [9, 11] of the magnetic subsystems $\mathcal{L}$ and $\mathcal{R}$, respectively. The coupling between the two magnetic subsystems is due to the layer off-diagonal projections $g_{1N}(\mathbf{k}_{\parallel}, z)$ and $g_{N1}(\mathbf{k}_{\parallel}, z)$ of the Green function of the spacer. The oscillatory behavior of interlayer coupling is due to the oscillatory behavior of these quasi one-dimensional spacer Green functions.

The structure constant matrix $S_{\mathbf{R},\mathbf{R}'}$ and the site-diagonal matrix associated with the potential function $P_{\mathbf{R}}(z)$ can be written in spin space as the following 2×2 supermatrices

$$P_{\mathbf{R}}(z) = \begin{pmatrix} P_{\mathbf{R}}^{\uparrow\uparrow}(z), & P_{\mathbf{R}}^{\uparrow\downarrow}(z) \\ P_{\mathbf{R}}^{\downarrow\uparrow}(z), & P_{\mathbf{R}}^{\downarrow\downarrow}(z) \end{pmatrix}, \quad S_{\mathbf{R},\mathbf{R}'} = \begin{pmatrix} S_{\mathbf{R},\mathbf{R}'}^{\uparrow\uparrow} & 0 \\ 0 & S_{\mathbf{R},\mathbf{R}'}^{\downarrow\downarrow} \end{pmatrix}, \quad S_{\mathbf{R},\mathbf{R}'}^{\uparrow\uparrow} = S_{\mathbf{R},\mathbf{R}'}^{\downarrow\downarrow}. \tag{13}$$

Two remarks might be now useful: (i) in the absence of spin-orbit coupling only the relative angle between the spin orientations in two different magnetic layers is important; (ii) the dipole-dipole magnetostatic interaction energy is often neglected as it becomes negligibly small for spacers containing three or more atomic layers.

The energy difference between the rotated and the ferromagnetic alignment is proportional to the quantity

$$\text{tr} \ln Z = \text{tr} \ln (1 - A_0 \, B) - \text{tr} \ln (1 - A_0 \, B_0), \tag{14}$$

where the matrices A_0 and B_0 are related to the ferromagnetic alignment and thus are diagonal in spin space

$$A_0 = \begin{pmatrix} A_0^{\uparrow} & 0 \\ 0 & A_0^{\downarrow} \end{pmatrix}, \quad B_0 = \begin{pmatrix} B_0^{\uparrow} & 0 \\ 0 & B_0^{\downarrow} \end{pmatrix}. \tag{15}$$

The particular form of the subblocks A_0^{σ} and B_0^{σ} ($\sigma = \uparrow, \downarrow$) is given by

$$A_0^{\sigma} = g_{N1}(\mathbf{k}_{\parallel}, z) \, \tau_1^{\sigma}(\mathbf{k}_{\parallel}, z) \, g_{1N}(\mathbf{k}_{\parallel}, z), \quad B_0^{\sigma} = \tau_N^{\sigma}(\mathbf{k}_{\parallel}, z). \tag{16}$$

The matrix B refers to an alignment in which the orientations of the spin in the two magnetic slabs are rotated relatively by an angle θ,

$$B = U(\theta) \, B_0 \, U^{\dagger}(\theta), \quad U(\theta) = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}, \tag{17}$$

where $U(\theta)$ is the rotation matrix for spin 1/2, $c = \cos(\theta/2)$, and $s = \sin(\theta/2)$. The final result is

$$\mathcal{E}_x(\theta) = \frac{1}{\pi N_{\parallel}} \sum_{\mathbf{k}_{\parallel}} \text{Im} \int_C f(z) \, \text{tr}_L \ln \left(1 - \frac{1 - \cos(\theta)}{2} M(\mathbf{k}_{\parallel}, z)\right) dz, \tag{18}$$

73

where
$$M = 1 - (1 - A_0^\uparrow B_0^\uparrow)^{-1} (1 - A_0^\uparrow B_0^\downarrow) (1 - A_0^\downarrow B_0^\downarrow)^{-1} (1 - A_0^\downarrow B_0^\uparrow). \tag{19}$$

By expanding the logarithm in (18) in powers of $[1 - \cos(\theta)]$, we find

$$\mathcal{E}_x(\theta) = \sum_{n=1}^{\infty} A_n \left[1 - \cos(\theta)\right]^n, \tag{20}$$

where A_n are the exchange coupling coefficients

$$A_n = -\frac{1}{2^n \, n \, \pi N_{\parallel}} \sum_{\mathbf{k}_{\parallel}} \operatorname{Im} \int_C f(z) \operatorname{tr}_L \left[M(\mathbf{k}_{\parallel}, z)\right]^n dz. \tag{21}$$

The expansion of $\mathcal{E}_x(\theta)$ for a small θ, i.e., when $1 - \cos(\theta)$ is a small parameter becomes particularly important when the spacer is a magnetic metal or for complicated geometries such as periodic multilayers. This approach is appropriate for infinitesimal rotations and is based on the local force theorem for magnetic systems, as originally developed for bulk magnetic crystals and alloys [47] (see also next subsection). In the case of Cu/Co/Cu/Co/Cu trilayers we have verified that the dependence of $\mathcal{E}_x(\theta)$ on the angle θ is approximated with a high accuracy by the first term $\mathcal{E}_x(\theta, N) \approx A_1(N) \left(1 - \cos(\theta)\right)$ of the expansion (20) not only for small angles θ but for any $\theta \leq \pi$.
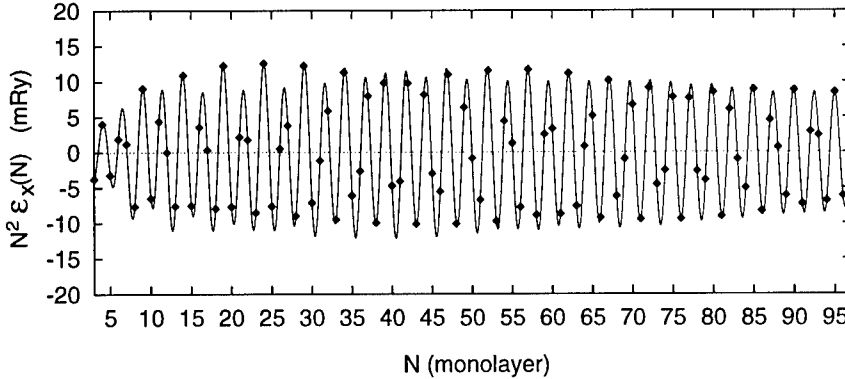
Figure 3: Exchange coupling $N^2 \mathcal{E}_x(N)$ as a function of the Cu spacer thickness N separating two semiinfinite fcc (001) Co slabs at $T = 0$ K.

In numerical calculations self-consistent potentials of bulk Cu and of a single Co(001) monolayer in bulk Cu are used also for interacting slabs (the so-called frozen potential approximation [42, 43, 44]). The energy and the Brillouin zone integrations need special care (see Refs. [42, 43, 44]).

As an illustration we present in Fig. 3 the dependence of the interlayer exchange coupling on the spacer thickness. Well-pronounced oscillations with the period 2.5 monolayers are found which decay with spacer thickness N as N^{-2}.

This theory can be generalized also to the case of disordered spacer, and/or disordered magnetic layers. The vertex-cancellation theorem [48] allows one to significantly simplify the study of various alloying effects in multilayers.

Heisenberg-type Hamiltonians

The parameters of the Heisenberg-type Hamiltonian (2) describing the elementary excitations in ferromagnets can be found using the method of infinitesimal rotations [47, 49] which is based on the force theorem and the Lloyd formula. The parameter $J_{\mathbf{RR'}}$ is proportional to the second variation of the band energy term E_{band} with respect to infinitesimal changes of the unit vectors $\delta e_{\mathbf{R}}$ and $\delta e_{\mathbf{R'}}$ that are represented by (infinitesimal) rotation matrices for spin $1/2$. In the case of a ferromagnetic reference state the first variations vanish and one finds

$$J_{\mathbf{RR'}} = \frac{1}{4\pi} \operatorname{Im} \int_{-\infty}^{\infty} f(z) \operatorname{tr}_L \left[\Delta_{\mathbf{R}}(z) \, g_{\mathbf{RR'}}^{\uparrow\uparrow}(z) \, \Delta_{\mathbf{R'}}(z) \, g_{\mathbf{R'R}}^{\downarrow\downarrow}(z) \right] dE \,, \qquad (22)$$

where $\Delta_{\mathbf{R}}(z) = P_{\mathbf{R}}^{\uparrow}(z) - P_{\mathbf{R}}^{\downarrow}(z)$ is the difference between the potential functions for spin up and spin down electrons, and tr_L denotes the trace over the orbital momentum $L = (\ell m)$. Note that the expression (22) is formally similar to that for unrenormalized pair interactions in the alloy case. This approach can be generalized to the case of disordered alloys as well as of surfaces. As an illustration we present in Table IV the values of $J_{\mathbf{RR'}}$ for $Fe_{85}V_{15}$ alloy.

Table IV: The parameters $J_{\mathbf{RR'}}$ (in mRy) of the Heisenberg-type Hamiltonian for Fe-Fe, Fe-V, and V-V pairs up to 9-th neighbors for $Fe_{85}V_{15}$ alloy as calculated by the method of infinitesimal rotations.

n	1	2	3	4	5	6	7	8	9
J_n^{Fe-Fe}	1.942	0.336	0.034	−0.013	−0.017	0.026	−0.003	0.005	−0.015
J_n^{Fe-V}	0.608	−0.035	0.023	0.044	−0.063	0.013	−0.002	−0.001	−0.011
J_n^{V-V}	−0.209	−0.043	−0.003	−0.001	0.016	0.001	0.000	0.000	0.003

Table V: Comparison of theoretical (mean-field) and experimental Curie temperatures T_C for ferromagnetic $3d$ metals and alloys. Also given are average magnetic moments μ.

	mean-field		experiment [39]	
	$\mu \, (\mu_B)$	T_C (K)	$\mu \, (\mu_B)$	T_C (K)
bcc Fe	2.26	1429	2.22	1043
fcc Co	1.61	1646	1.75	1388
fcc Ni	0.60	392	0.62	627
bcc $Fe_{75}Co_{25}$	2.34	2138	2.45	1220
bcc $Fe_{50}Co_{50}$	2.16	2020	2.36	1255
bcc $Fe_{25}Co_{75}$	1.94	1774	2.14	1170
bcc $Fe_{85}V_{15}$	1.70	1601	1.72	1070
bcc $Fe_{75}V_{25}$	1.37	1271	1.41	1050
bcc $Fe_{90}Cr_{10}$	1.94	1589	1.97	1020
bcc $Fe_{85}Cr_{15}$	1.81	1432	1.84	970
bcc $Fe_{50}Cr_{50}$	1.02	545	1.00	600

The parameters $J_{\mathbf{RR'}}$ can be used in statistical mechanical studies of the ferromagnetic system, e.g. to determine the Curie temperature T_C or the spin stiffness [47, 49]. According

to a simple mean-field theory

$$T_C = \frac{2J_0}{3k_B}, \quad J_0 = \sum_{\mathbf{R}'(\neq\mathbf{R})} J_{\mathbf{R}\mathbf{R}'}, \tag{23}$$

where the quantity J_0 can also be calculated directly without necessity to sum $J_{\mathbf{R}\mathbf{R}'}$ over all coordination spheres [47, 49]. A few examples are given in Table V. As noted earlier, the mean-field theory overestimates the values of T_C. It was shown [50] that a more accurate treatment, based on Monte Carlo simulations, leads to a substantial reduction of T_C.

CONCLUSIONS

We have shown that the energetics of a wide spectrum of systems can be studied within the ab initio framework by using a unified and highly efficient formalism based on the force theorem and the Lloyd formula. Within this formalism the total energies of various configurations of the system are mapped onto simple effective Hamiltonians from which the physical properties can be extracted by employing methods of statistical mechanics. We have shown that ab initio tight-binding techniques that can be formulated in the Green function language, such as the screened linear muffin-tin orbital method, are particularly suitable for this purpose. We have presented here several applications of this approach to systems as diverse as bulk metallic alloys, their surfaces and interfaces, two-dimensional alloys, magnetic multilayers, and ferromagnets.

ACKNOWLEDGMENTS

REFERENCES

1. I. Turek, V. Drchal, J. Kudrnovský, M. Šob, and P. Weinberger,
   Electronic Structure of Disordered Alloys, Surfaces and Interfaces, (Kluwer Academic, Boston, 1997).
2. V. Drchal, J. Kudrnovský, and I. Turek, Computer Phys. Commun. **97**, 111 (1996).
3. J. Kudrnovský, I. Turek, V. Drchal, and M. Šob, in Stability of Materials, edited by A. Gonis, P. Turchi, and J. Kudrnovský (Plenum, New York, 1996), pp. 237-264.
4. O.K. Andersen, Phys. Rev. B **12**, 3060 (1975).
5. H.L. Skriver, The LMTO Method, (Springer, Berlin, 1984).
6. O.K. Andersen and O. Jepsen, Phys. Rev. Lett. **53**, 2571 (1984).
7. O.K. Andersen, Z. Pawlowska, and O. Jepsen, Phys. Rev. B **34**, 5253 (1986).
8. J. Kudrnovský and V. Drchal, Phys. Rev. B **41**, 7515 (1990).
9. B. Wenzien, J. Kudrnovský, V. Drchal, and M. Šob, J. Phys.: Condens. Matter **1**, 9893 (1989).
10. H.L. Skriver and N.M. Rosengaard, Phys. Rev. B **43**, 9538 (1991).

11. J. Kudrnovský, I. Turek, V. Drchal, P. Weinberger, N. E. Christensen, and S. K. Bose, Phys. Rev. B **46**, 4222 (1992).

12. J. Kudrnovský, I. Turek, V. Drchal, P. Weinberger, S. K. Bose, and A. Pasturel, Phys. Rev. B **47**, 16525 (1993).

13. A.B. Shick, V. Drchal, J. Kudrnovský, and P. Weinberger: Phys. Rev. B **54**, 1610 (1996).

14. V. Drchal, J. Kudrnovský, and P. Weinberger, Phys. Rev. B **50**, 7903 (1994).

15. J.W. Connolly and A.R. Williams, Phys. Rev. B **27**, 5169 (1983).

16. A.R. Mackintosh and O.K. Andersen, Chapter 5.3 in Electrons at the Fermi Surface, edited by M. Springford (Cambridge University Press, Cambridge, England, 1980).

17. P. Lloyd and P.V. Smith, Adv. Phys. **21**, 69 (1972).

18. F. Ducastelle and F. Gautier, J. Phys. F: Metal Phys. **6**, 2039 (1976).

19. F. Ducastelle, Order and Phase Stability, (North Holland, Amsterdam, 1991).

20. P.E.A. Turchi, Chapter 2 in Intermetallic Compounds, edited by J.H. Westbrook and R.L. Fischer (Wiley, New York, 1994).

21. V. Drchal, J. Kudrnovský, L. Udvardi, P. Weinberger, and A. Pasturel, Phys. Rev. B **45**, 14328 (1992).

22. A. Pasturel, V. Drchal, J. Kudrnovský, and P. Weinberger, Phys. Rev. B **48**, 2704 (1993).

23. F. Ducastelle, J. Phys. C: Solid State Phys. **8**, 3297 (1975).

24. A. Gonis, P.E.A. Turchi, J. Kudrnovský, V. Drchal, and I. Turek, J. Phys.: Condensed Matter **8**, 7869, (1996); ibid. 7883, (1996).

25. S.K. Bose, V. Drchal, J. Kudrnovský, O. Jepsen, and O.K. Andersen, Phys. Rev. B **55**, 8184 (1997).

26. V. Drchal, J. Kudrnovský, A. Pasturel, I. Turek, and P. Weinberger, Phys. Rev. B **54**, 8202 (1996).

27. K. Binder and D.W. Heermann, Monte Carlo Simulation in Statistical Physics, (Springer, Berlin, 1994).

28. J. Kudrnovský, I. Turek, A. Pasturel, R. Tétot, V. Drchal, and P. Weinberger, Phys. Rev. B **50**, 9603 (1994).

29. R. Monnier, Phil. Mag. B **75**, 67 (1997).

30. B. Legrand, G. Tréglia, and F. Ducastelle, Phys. Rev. B **41**, 4422 (1990); G. Tréglia, B. Legrand, and F. Ducastelle, Europhys. Lett. **7**, 575 (1988).

31. T. Schulthess, R. Monnier, and S. Crampin, Phys. Rev. B **50**, 18564 (1994).

32. V. Drchal, J. Kudrnovský, and I. Turek, in Stability of Materials, edited by A. Gonis, P. Turchi, and J. Kudrnovský (Plenum, New York, 1996), pp. 355-360.

33. V. Drchal, A. Pasturel, J. Kudrnovský, A. Gonis, and P. Turchi, in Properties of Complex Inorganic Solids, edited by A. Gonis, A. Meike, and P.E.A. Turchi, (Plenum, New York, 1997), pp. 133-138.

34. J. Florencio, D.M. Ren, and T.T. Tsong, Surf. Sci. **345**, L29 (1996).

35. B. Legrand and G. Tréglia, Surf. Sci. **236**, 398 (1990).

36. J. Kudrnovský, S.K. Bose, and V. Drchal, Phys. Rev. Lett. **69**, 308 (1992).

37. R. Tétot, J. Kudrnovský, A. Pasturel, V. Drchal, and P. Weinberger, Phys. Rev. B **51**, 17910 (1995).

38. <u>Binary Alloy Phase Diagrams</u>, 2nd edition, editor-in-chief T.B. Massalski, (ASM International, Materials Park, Ohio, 1990).

39. <u>Landolt-Börnstein-New Series</u> III, Vol. 19a, edited by O. Madelung, (Springer, Berlin, 1992).

40. A.G. Khachaturyan, <u>Theory of Structural Transformations in Solids</u>, (Wiley, New York, 1983), Chap. 3.7.

41. A. Bieber and F. Gautier, J. Magn. Magn. Mater. **97**, 293 (1991).

42. P. Lang, L. Nordström, R. Zeller, and P.H. Dederichs, Phys. Rev. Lett. **71**, 1927 (1993).

43. J. Kudrnovský, V. Drchal, I. Turek, and P. Weinberger, Phys. Rev. B **50**, 16105 (1994).

44. P. Lang, L. Nordström, K. Wildberger, R. Zeller, P.H. Dederichs, and T. Hoshino, Phys. Rev. B **53**, 9092 (1996).

45. K.B. Hathaway, <u>Ultrathin Magnetic Structures II</u>, edited by B. Heinrich and J.A.C. Bland (Springer, Berlin, 1994), 45.

46. J. d' Albuquerque e Castro, M.S. Ferreira, and R.B. Muniz, Phys. Rev. B **49**, 16062 (1994).

47. A.I. Liechtenstein, M.I. Katsnelson, V.P. Antropov, and V.A. Gubanov, J. Magn. Magn. Mater. **67**, 65 (1987).

48. P. Bruno, J. Kudrnovský, V. Drchal, and I. Turek, Phys. Rev. Lett. **76**, 4254 (1996).

49. V.A. Gubanov, A.I. Liechtenstein, and A.V. Postnikov, <u>Magnetism and the Electronic Structure of Crystals</u>, (Springer, Berlin, 1992).

50. N.M. Rosengaard and B. Johansson, Phys. Rev. B **55**, 14975 (1997).

ELECTRONIC STRUCTURE AND ATOMIC CONFIGURATION OF EXTENDED DEFECTS IN METALS BY FIRST-PRINCIPLES AND SEMIEMPIRICAL TB-LMTO METHODS

M. Šob[*,**], I. Turek[*], and V. Vitek[**]
[*]Institute of Physics of Materials, Academy of Sciences of the Czech Republic,
Žižkova 22, CZ-616 62 Brno, Czech Republic, mojmir@ipm.cz
[**]Department of Materials Science and Engineering, University of Pennsylvania,
3231 Walnut St., Philadelphia, PA 19104-6272, U. S. A., vitek@sol1.lrsm.upenn.edu

ABSTRACT

We present two tight-binding linear muffin-tin orbitals (TB-LMTO) techniques for electronic structure calculations of extended defects (such as grain boundaries, interphase interfaces, surface layers etc.) in metals. The first is based on the first-principles self-consistent surface Green's function approach within the atomic-sphere approximation (ASA) utilizing two-dimensional periodicity in the layers parallel to the interface. In the second approach the Hamiltonian is constructed within the TB-LMTO-ASA as well, but semiempirical terms are employed to characterize the repulsive part of the interaction and the effect of electrons in interstitial space. While the adjustable parameters have only been fitted to the properties of ideal ground state structure, the semiempirical approach describes correctly the structural energy differences, phonon frequencies etc. Two examples are presented: the electronic structure of the $\Sigma = 5(210)/[001]$ tilt grain boundary in tungsten is determined and the sensitivity of 4d magnetic moments in thin films to local environment is discussed. A comparison of the semiempirical TB-LMTO-ASA with the first-principles full-potential LMTO results is performed along the trigonal deformation path connecting the bcc, simple cubic and fcc structures and the applicability of the semiempirical approach for simulating atomic structure of extended defects is assessed.

1. INTRODUCTION

In the tight-binding (TB) approach, the basis functions used for the expansion of one-electron states are atomic-like orbitals, displaying the same angular momentum as orbitals associated with a free atom (i.e. s, p, d etc. corresponding to angular momentum $\ell = 0, 1, 2$ etc.), but with radial dependencies that can differ significantly from free atomic orbitals. Once regarded as suitable only for a very restricted class of materials (i.e. those where the overlap between atomic basis functions on neighbouring atoms is small), the TB procedure has been shown to give very satisfactory results even e.g. for nearly free electron systems such as alkali metals [1]. The advantage over the methods based on plane waves is that a relatively small basis set may be employed; a disadvantage is that extensive computation of complicated integrals may be needed for quantitative results in first-principles applications.

However, if some information is known from experiment or accurate computations about the electronic structure (ES) of a solid, then one can often fit these integrals and interpolate, for example, the energy bands. In this case the matrix elements in the secular equation are directly fitted and the orbitals need not be (and usually are not) constructed. However, the charge density distribution is then not calculated and electronic structure in the TB calculations cannot be determined self-consistently in the sense of the density functional theory.

In this semiempirical representation, the TB method is not a rigorous, first-principles theory any more. However, it is very attractive from the point of view of the conceptual transparency and computational simplicity. Very often an insight into the bonding processes

and a deeper understanding of the trends in properties from one system to another may be gained from some empirical or model Hamiltonian which is usually much simpler than that employed in ab initio approaches. Thus the TB method is a very effective tool to determine trends when some parameters (e.g. electronic concentration in alloys) are varied, particularly when dealing with complex systems far from ideal simple crystals (e.g. surfaces, defects, incoherent interfaces, amorphous structures). The large-scale first-principles ES calculations are then essential for both giving the credibility to the results obtained from semiempirical model Hamiltonians and to supply necessary input parameters.

The description of the TB method may be found e.g. in [2–5], a discussion of the perspectives and generalized functionals for the TB method is performed in [6, 7], and various linear-scaling TB approaches are compared in [8].

The TB method has also been reformulated within the screened LMTO [9–12] and KKR [13] formalism. At present, the third generation of the LMTO method is being developed [14, 15].

In this contribution, we will show how to treat atomic configuration and electronic structure of extended defects such as grain boundaries, interphase interfaces, surface layers etc., both within a fully first-principles and semiempirical TB-LMTO scheme.

2. FIRST-PRINCIPLES TB-LMTO THEORY

The tight-binding linear muffin-tin orbital (TB-LMTO) method was originally formulated by transforming the basis of the conventional linear muffin-tin orbitals into a new basis of the tight-binding linear muffin-tin orbitals [9–12]. The physical motivation for this transformation came from the idea of screened electrostatic multipole fields. In the screened LMTO theory, real constants $\alpha_{\mathbf{R}L}$ are introduced, defined for each $\mathbf{R}L$-index, traditionally called screening constants. The screening matrix $\alpha_{\mathbf{R}L,\mathbf{R}'L'} = \alpha_{\mathbf{R}L}\delta_{\mathbf{R}\mathbf{R}'}\delta_{LL'}$ characterizes the corresponding LMTO representation.

Most of the calculations using the LMTO method have been performed within so-called atomic-sphere approximation (ASA) [16–19]. This approximation is as follows: (i) spherically symmetric potentials inside slightly overlapping, space-filling atomic (Wigner-Seitz) spheres centered at the individual nuclei are used, and (ii) the electronic kinetic energy outside the Wigner-Seitz spheres is neglected. Therefore, ASA removes the inconvenient interstitial region and replaces integrals over the whole space by a sum over Wigner-Seitz spheres which may be considered as approximations to the true Wigner-Seitz cells. The overlap of the atomic spheres is neglected. The impact of these approximations is discussed in more details in Sec. 5.

The LMTO-ASA calculations use typically 9 (*spd*) or 16 (*spdf*) muffin-tin orbitals per non-equivalent atomic site. They have been very successful in calculations of electronic, co-hesive and magnetic properties for a large number of materials with close-packed structures, where the error due to overlapping atomic spheres is not very large. The LMTO method has also been used for very large supercells as well as for Green's function calculations for disordered alloys, surfaces and interfaces and localized and extended defects (for a recent review see [20] and references therein).

The LMTO-ASA overlap and Hamiltonian matrix elements in the α-representation are

$$O^\alpha = (1 + h^\alpha o^\alpha)(o^\alpha h^\alpha + 1) + h^\alpha p h^\alpha,$$
$$H^\alpha = h^\alpha(1 + o^\alpha h^\alpha) + (1 + h^\alpha o^\alpha)E_\nu(o^\alpha h^\alpha + 1) + h^\alpha E_\nu p h^\alpha, \tag{1}$$

where

$$h^\alpha_{\mathbf{R}L,\mathbf{R}'L'} = \xi^\alpha_{\mathbf{R}\ell}\left(C_{\mathbf{R}\ell} - E_{\nu,\mathbf{R}\ell}\right)\delta_{\mathbf{R}L,\mathbf{R}'L'}$$
$$+ \xi^\alpha_{\mathbf{R}\ell}\sqrt{\Delta_{\mathbf{R}\ell}}\, S^\alpha_{\mathbf{R}L,\mathbf{R}'L'}\,\xi^\alpha_{\mathbf{R}'\ell'}\sqrt{\Delta_{\mathbf{R}'\ell'}}. \tag{2}$$

The quantities $o^\alpha_{\mathbf{R}\ell}$ and $\xi^\alpha_{\mathbf{R}\ell}$ are explicitly given by

$$o^\alpha_{\mathbf{R}\ell} = \frac{\alpha_{\mathbf{R}\ell} - \gamma_{\mathbf{R}\ell}}{\xi^\alpha_{\mathbf{R}\ell} \Delta_{\mathbf{R}\ell}},$$

$$\xi^\alpha_{\mathbf{R}\ell} = 1 + \frac{(\alpha_{\mathbf{R}\ell} - \gamma_{\mathbf{R}\ell})(C_{\mathbf{R}\ell} - E_{\nu,\mathbf{R}\ell})}{\Delta_{\mathbf{R}\ell}}. \tag{3}$$

Further, C, Δ, γ and p are the diagonal matrices of LMTO potential parameters character-
izing scattering properties of atoms in the solid and E_ν are the energies chosen usually at
the center of interest in the neighbourhood of which one performs linearization [10, 12, 19].
They are dependent only on the orbital quantum number ℓ, not on the magnetic quantum
number m. The parameters $C_{\mathbf{R}\ell}$ determine essentially the positions of the $\mathbf{R}\ell$ band, whereas
$\Delta_{\mathbf{R}\ell}$ determines the width and hybridization strength of the $\mathbf{R}\ell$ band. For example, it was
shown that the width of a transition metal d band is approximately 25 Δ_d [10].

The matrix $S^\alpha_{\mathbf{R}LR'L'}$ is so-called structure constant matrix. It is determined solely by the
positions of atoms, i.e. by the geometry. Therefore, the Hamiltonian (1) can be factorized
into the crystal potential part, given by potential parameters, and the structure part,
determined by the structure constant matrix. When seeking the optimal screening [10, 12],
it has been assumed that there is one value of the screening parameter $\alpha_{\mathbf{R}L}$ for all angular
momentum indices $L \equiv (\ell m)$ with the same orbital quantum number ℓ, independently of
the site $\mathbf{R}$, i.e.

$$\alpha_{\mathbf{R}L,\mathbf{R}'L'} = \alpha_\ell \delta_{\mathbf{R}\mathbf{R}'}\delta_{LL'}. \tag{4}$$

One has then, therefore, one α_s for all s-orbitals, one α_p for all p-orbitals etc. The values of
these optimal screening parameters for various types of screening are as follows:

Table 1: Screening constants α for best overall screening in a number of structures.

screening type	ℓ_{max}	ref.	α_s	α_p	α_d	α_f
s	0	[12]	0.2143			
sp	1	[12]	0.2872	0.02582		
spd	2	[9, 10]	0.3485	0.05303	0.01071	
spdf	3	[21]	0.3851	0.07321	0.02248	0.00607

Owing to the short-range of the screened structure constants and the LMTO's, the
screened LMTO representation with α's given by Table 1 is called tight-binding (TB) LMTO
representation and the corresponding matrix S^α is the TB structure constant matrix. The
hopping integrals in the LMTO-ASA method are then given by

$$\beta_{\mathbf{R}L,\mathbf{R}'L'} = \xi^\alpha_{\mathbf{R}\ell} \sqrt{\Delta_{\mathbf{R}\ell}} \ S^\alpha_{\mathbf{R}L,\mathbf{R}'L'} \ \xi^\alpha_{\mathbf{R}'\ell'} \sqrt{\Delta_{\mathbf{R}'\ell'}}. \tag{5}$$

To find the electron energies and the corresponding wavefunctions, one should solve the
eigenvalue problem for the matrix $H^\alpha - EO^\alpha$ with a non-unity overlap matrix. However,
it is possible to transform the whole secular matrix to a nearly orthogonal representation.
Assuming that the matrix $(1 + o^\alpha h^\alpha)$ may be inverted, it follows from the structure of the
Hamiltonian and overlap matrices (1) that if we transform to a new basis defined as

$$|\chi^\gamma\rangle \equiv |\chi^\alpha\rangle)(1 + o^\alpha h^\alpha)^{-1} \tag{6}$$

the Hamiltonian and the overlap matrix (1) are given in this basis by

$$H^\gamma = E_\nu + h^\gamma + h^\gamma E_\nu p h^\gamma, \qquad O^\gamma = 1 + h^\gamma p h^\gamma \tag{7}$$

with

$$h^\gamma \equiv h^\alpha (1 + o^\alpha h^\alpha)^{-1}. \tag{8}$$

Thus, the set of orbitals $|\chi^\gamma\rangle$ is orthogonal to the first order in h^γ.

In many LMTO-ASA calculations the parameter p has often been neglected and the orthogonal Hamiltonian H^γ contains in this case only two-center terms ($o^\gamma = 0$ in the orthogonal representation). On the other hand, the range of the corresponding structure constant matrix $S^\gamma_{\mathbf{R}L,\mathbf{R}'L'}$ is much bigger (up to 7th neighbours [10]), so that this representation cannot be considered any more as tight-binding.

3. GREEN'S FUNCTION WITHIN THE TB-LMTO-ASA

Conventional band-structure methods are of limited use in studies of more complex systems like disordered alloys, imperfect surfaces, interfaces, multilayered structures etc. because they require an excessive number of atoms per repeat cell. In addition, they are not able to account fully for substitutional disorder and the true semi-infinite geometry of surfaces. Such problems can be solved more appropriately by Green's function techniques and multiple scattering formalism.

The TB-LMTO-ASA technique is highly convenient for Green's function calculations on perturbed crystals. As shown in [11, 20], the physical Green's function $G(z)$ within the TB-LMTO-ASA formalism is identical to the resolvent of the orthogonal LMTO Hamiltonian H^γ:

$$G_{\mathbf{R}L,\mathbf{R}'L'}(z) = [(z - H^\gamma)^{-1}]_{\mathbf{R}L,\mathbf{R}'L'} \tag{9}$$

It may be expressed as

$$G_{\mathbf{R}L,\mathbf{R}'L'}(z) = \lambda^\alpha_{\mathbf{R}L}(z)\,\delta_{\mathbf{R}\mathbf{R}'}\,\delta_{LL'} + \mu^\alpha_{\mathbf{R}L}(z)\,g^\alpha_{\mathbf{R}L,\mathbf{R}'L'}(z)\,\mu^\alpha_{\mathbf{R}'L'}(z), \tag{10}$$

where the auxiliary Green's function $g^\alpha(z)$ is defined by

$$g^\alpha_{\mathbf{R}L,\mathbf{R}'L'}(z) = [(P^\alpha(z) - S^\alpha)^{-1}]_{\mathbf{R}L,\mathbf{R}'L'}. \tag{11}$$

Here $P^\alpha(z)$ is the potential function characterizing the scattering properties of atoms in the LMTO representation α; the site-diagonal matrices $\lambda^\alpha(z)$ and $\mu^\alpha(z)$ are expressed by means of this potential function and their derivatives, i.e. in terms of potential parameters [20].

The most important advantage of the Green's function formulation is that the physical Green's function $G(z)$ is invariant with respect to screening constants α. It is therefore possible to choose such LMTO representation that is most suitable for a particular physical problem. For example, when treating the ES of disordered alloys, the configurational averaging of the Green's function is performed in a non-random LMTO representation so that the Green's function does not exhibit off-diagonal randomness. Here the TB representation which is also non-random is usually employed [20].

Once the Green's function is found, electronic charge density, local density of states etc. may easily be calculated. For example, the projected local density of states for an atom at $\mathbf{R}$ is given by

$$n_{\mathbf{R}L}(E) = -\frac{1}{\pi}\,\mathrm{Im}\,G_{\mathbf{R}L,\mathbf{R}L}(E + i0). \tag{12}$$

4. GREEN'S FUNCTION METHOD FOR PLANAR DEFECTS

A system with a planar extended defect may be considered as composed of two unperturbed semi-infinite bulk systems sandwiching the perturbed (intermediate) region. (In

82

the case of surfaces, one of these bulks is vacuum.) The perturbation may extend over several atomic layers. When studying the ES of the system, we take advantage of the two-dimensional periodicity in the layers parallel to the interface and the short range of the hopping integrals (structure constants) in the TB-LMTO method [10, 12, 20]. The former enables us to use the Bloch theorem in two dimensions and to introduce the Bloch wave vector $\mathbf{k}_{\parallel}$, the latter leads to a notion of the principal layers [20, 22]. Each principal layer contains such (finite) number of neighboring atomic layers that only the nearest principal layers interact. The whole system may then be considered as a sequence of principal layers, both in the semi-infinite bulks and in the intermediate region.

The TB-LMTO-ASA Green's function of the whole system may be expressed as

$$G_{p\mathbf{B}L,p'\mathbf{B}'L'}(\mathbf{k}_{\parallel}, z) = \lambda_{p\mathbf{B}L}^{\alpha}(z)\, \delta_{p\mathbf{B}L,p'\mathbf{B}'L'} + \mu_{p\mathbf{B}L}^{\alpha}(z)\, g_{p\mathbf{B}L,p'\mathbf{B}'L'}^{\alpha}(\mathbf{k}_{\parallel}, z)\, \mu_{p'\mathbf{B}'L'}^{\alpha}(z), \qquad (13)$$

where p and p' label the principal layers, $\mathbf{B}$ and $\mathbf{B}'$ denote non-equivalent sites in the principal layers, L and L' are usual angular momentum indices, λ^{α} and μ^{α} are site-diagonal matrices and z is the complex energy. The auxiliary Green's function is given by

$$(g^{\alpha})_{p\mathbf{B}L,p'\mathbf{B}'L'}^{-1}(\mathbf{k}_{\parallel}, z) = P_{p\mathbf{B}L}^{\alpha}(z)\, \delta_{p\mathbf{B}L,p'\mathbf{B}'L'} - S_{p\mathbf{B}L,p'\mathbf{B}'L'}^{\alpha}(\mathbf{k}_{\parallel}), \qquad (14)$$

where $P^{\alpha}(z)$ is the site-diagonal potential function matrix and $S^{\alpha}(\mathbf{k}_{\parallel})$ is the matrix of TB-LMTO structure constants [10, 12, 20].

The main problem now is to invert the infinite matrix $(g^{\alpha})^{-1}$ in eq. (14), which is block-tridiagonal in the principal layer indices p, p' (Fig. 1; in this particular example, principal layers labelled by $p = 1\text{–}4$ belong to the intermediate region).

$$\mathbf{g}^{-1}(\mathbf{k}_{\parallel}, z) =$$

	-1	0	1	2	3	4	5	6
-1	$\mathbf{P}_{-1} - \mathbf{S}_{-1-1}$	$-\mathbf{S}_{-10}$						
0	$-\mathbf{S}_{0-1}$	$\mathbf{P}_0 - \mathbf{S}_{00}$	$-\mathbf{S}_{01}$					
1		$-\mathbf{S}_{10}$	$\mathbf{P}_1 - \mathbf{S}_{11}$	$-\mathbf{S}_{12}$				
2			$-\mathbf{S}_{21}$	$\mathbf{P}_2 - \mathbf{S}_{22}$	$-\mathbf{S}_{23}$			
3				$-\mathbf{S}_{32}$	$\mathbf{P}_3 - \mathbf{S}_{33}$	$-\mathbf{S}_{34}$		
4					$-\mathbf{S}_{43}$	$\mathbf{P}_4 - \mathbf{S}_{44}$	$-\mathbf{S}_{45}$	
5						$-\mathbf{S}_{54}$	$\mathbf{P}_5 - \mathbf{S}_{55}$	$-\mathbf{S}_{56}$
6							$-\mathbf{S}_{65}$	$\mathbf{P}_6 - \mathbf{S}_{66}$
	bulk I		intermediate region				bulk II	

Fig. 1. The structure of the infinite matrix $(g^{\alpha})^{-1}(\mathbf{k}_{\parallel}, z)$. This matrix is block-tridiagonal in the principal layer indices p, p'. The intermediate region is represented by principal layers 1-4. The superscript α denoting the LMTO representation has been dropped.

For this purpose we use the TB-LMTO surface Green's function technique [23, 24, 25, 20]. The surface Green's functions g^{bI} and g^{bII} for the bulk I and bulk II (see Fig. 1) may be

found from the removal invariance (i.e. from the fact that removing a finite number of layers at the end of a homogeneous semi-infinite stacking of layers recovers the original semi-infinite system [5, 22, 23]). Using the relations (with the superscript α omitted)

$$\begin{aligned}
P_0 &= P_{-1} = P_{-2} = ... \equiv P^{bI} \\
P_5 &= P_6 = P_7 = ... \equiv P^{bII},
\end{aligned} \qquad (15)$$

we obtain

$$\begin{aligned}
g^{bI} &= [P^{bI} - S_{00} - S_{0-1}g^{bI}S_{-10}]^{-1} \\
g^{bII} &= [P^{bII} - S_{55} - S_{56}g^{bII}S_{65}]^{-1}.
\end{aligned} \qquad (16)$$

The quantities $\Gamma^{bI} = S_{10}g^{bI}S_{01}$ and $\Gamma^{bII} = S_{45}g^{bI}S_{54}$ are "embedding potentials" describing the influence of two semi-infinite systems adjacent to the intermediate region. If we are interested only in the blocks of g within the intermediate region, we may then invert the matrix (the superscript α is also omitted here)

$$g^{-1}(\mathbf{k}_{\|}, z) =$$

p, p'	1	2	3	4
1	$P_1 - S_{11}$ $-S_{10}g^{bI}S_{01}$	$-S_{12}$		
2	$-S_{21}$	$P_2 - S_{22}$	$-S_{23}$	
3		$-S_{32}$	$P_3 - S_{33}$	$-S_{34}$
4			$-S_{43}$	$P_4 - S_{44}$ $-S_{45}g^{bII}S_{54}$

This is a matrix of finite dimensions. After performing the inversion, the physical Green's function may be calculated according to eq. (13). It yields the charge densities, densities of states (12) and other information about the ES of the system [5, 20].

It is important to note that the potential functions for the bulk systems sandwiching the perturbed region are usually obtained from separate standard selfconsistent calculations for the bulk. In the defect calculation, only intermediate region has to be treated selfconsistently.

As an example, we present the results for the $\Sigma = 5(210)/[001]$ tilt grain boundary (GB) in tungsten. To obtain the atomic structure of this GB, the relaxations were performed using the potentials of the Finnis-Sinclair type [26]. The electronic structure was calculated using an intermediate region consisting of 5 principal layers. Each principal layer contained 4 atomic layers (we included interactions till the second nearest neighbors).

The atomic configuration of the GB obtained using the Finnis-Sinclair type potentials is shown in Fig. 2. Here atoms numbered 9-14 lie in the most perturbed region.

It follows from the analysis of the local densities of states (LDOS) on individual atoms that away from the "core" of the grain boundary (atoms 1–7, 17–20) the LDOS is very similar to that of the bulk. However, in the most perturbed GB region we may observe smoothing of the DOS and filling of the pseudogap between bonding and antibonding d-states (Fig. 3). This is the consequence of a lower symmetry of the local neighborhood of atoms in the GB. Similar situation arises in amorphous materials.

Let us note that we have not detected any narrowing of the occupied portion of the d-band due to the reduced coordination, as found e.g. in [27]. Namely, in our relaxed GB structure, the coordination is reduced only slightly or not at all. Due to the filling of the

pseudogap at the Fermi energy, the bonding d-like peak in the LDOS is broadened for many atoms. On some sites, the whole peak is shifted by the local Madelung potential.
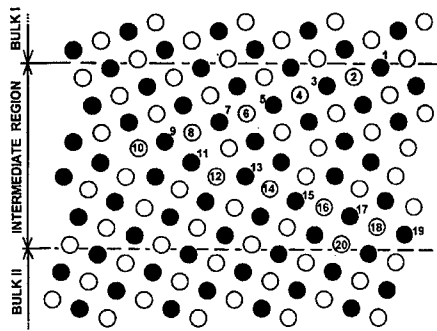
Fig. 2. Atomic configuration of the $\Sigma=5(210)/[001]$ tilt grain boundary in tungsten. Open circles represent atoms displaced by $\pm a[0, 0, 1/2]$ with respect to the atoms denoted by full circles (a is the lattice constant).
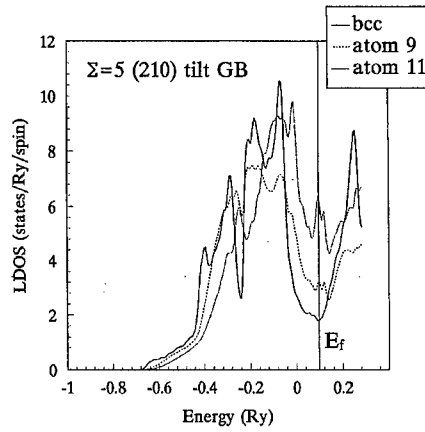
Fig. 3. Local densities of states at atoms 9 and 11 in the $\Sigma = 5(210)/[001]$ tilt grain boundary in tungsten (see Fig. 2) compared with the density of states for the ideal bcc tungsten.

As another example we present the results of our study of sensitivity of 4d-magnetic moments to local environment, in particular to structural imperfections.
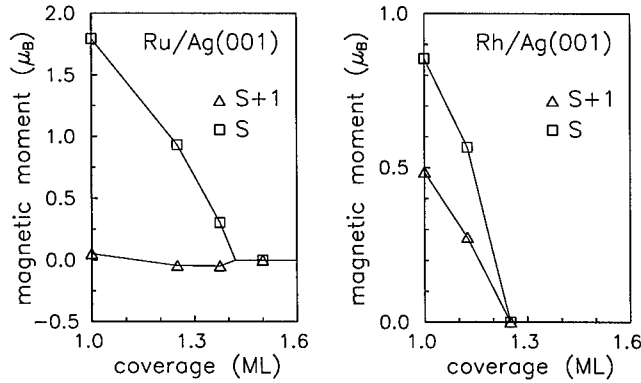
Fig. 4. Local magnetic moments in Ru (left panel) and Rh (right panel) overlayers on a Ag(001) substrate as a function of coverage. Squares and triangles refer to moments in the complete (S) and incomplete (S+1) transition-metal layers, respectively.

Epitaxial transition-metal overlayers of monolayer thickness on non-magnetic substrates like Cu, Ag, Au, Pd, Pt can be considered as realizations of two-dimensional itinerant magnetic systems. Recent systematic FLAPW calculations [28, 29] predicted spontaneous magnetic order for $4d$ and $5d$ transition-metal monolayers like Ru, Rh, Ir on Ag(001) and Au(001) substrates. However, despite the unambiguous theoretical results predicting the magnetism of these monolayers, experiments using the surface magneto-optic Kerr effect failed to confirm this phenomenon [30, 31]. The most probable reason for this discrepancy seems to be structural imperfections in the samples which are far off from perfect monolayers considered in the calculations. Ample experimental evidence demonstrates that islanding, surface roughness or interdiffusion cannot be neglected in these systems.

Combining the surface Green's function approach described above with the coherent potential approximation [20], we have studied the influence of two types of structural imperfections on the ferromagnetism of Ru and Rh monolayers on a Ag(001) substrate using a model of two-dimensional random alloys formed in several surface layers [32]. The first case corresponds to a non-integer coverage of the substrate with Ru or Rh varying between 1 and 2 monolayers, which was simulated by surface layers of composition $R_x Vac_{1-x}/R/Ag(001)$, $0 \leq x \leq 1$, where R stands for Ru or Rh and Vac denotes a surface vacancy. This model describes a continuous formation of the second R layer on top of the first perfect R layer on the Ag(001) face. The dependence of the local moments of R atoms with respect to the coverage is shown in Fig. 4.
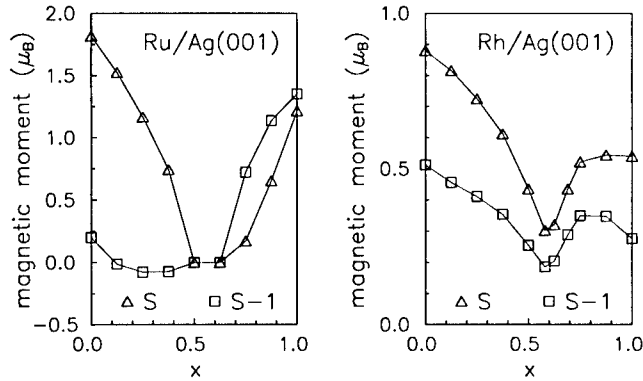
Fig. 5. Influence of interdiffusion on the local magnetic moments in Ru (left panel) and Rh (right panel) overlayers on a Ag(001) substrate. Triangles refer to the moments in the top surface (S) layer, squares refer to the first subsurface (S-1) layer. For details, see text.

In the second case, the importance of interdiffusion of Ru and Rh films of 1 monolayer coverage with the Ag(001) substrate was studied by simulating a layer sequence $R_{1-x}Ag_x/R_xAg_{1-x}/Ag(001)$, $0 \leq x \leq 1$, with all atoms occupying the positions of the ideal bulk Ag lattice. This model describes a continuous transition of the R layer from the surface ($x = 0$) to the first subsurface ($x = 1$) position. The corresponding concentration dependence of the local moments is presented in Fig. 5.

As can be seen from Figs. 4 and 5, despite the relatively large local moments in the perfect monolayer ($M_{Ru} = 1.8\,\mu_B$, $M_{Rh} = 0.9\,\mu_B$), the local moments of Ru and Rh atoms are highly sensitive to both types of structural imperfections. In the first model, two-dimensional ferromagnetism does not survive for coverages higher than 1.5 monolayer (Fig. 4) which

is in reasonably good agreement with FLAPW calculations for perfect double layers [33] yielding nearly negligible local moments for $4d$ and $5d$ transition-metal double layers on a Ag(001) substrate. An even more complicated behavior evolves from the second model of structural imperfections where a perfect R monolayer in the surface ($x = 0$) as well as in the first subsurface ($x = 1$) layer carries a sizable local moment while ferromagnetic behavior of monolayers strongly perturbed by interdiffusion (for intermediate values of x) is substantially reduced (Fig. 5). This leads to a magnetic gap for $0.5 \leq x \leq 0.6$ in the Ru case and to a deep minimum in both local moments near $x = 0.6$ in the Rh case.

The rapid decrease of the local moments with increasing coverage (Fig. 4) and interdiffusion (Fig. 5) can be understood from the shape of the local DOSs of R atoms; a more detailed discussion may be found in [32].

5. LIMITATIONS OF THE FIRST-PRINCIPLES LMTO-ASA THEORY

Due to presence of overlapping spheres in the LMTO-ASA formalism, it is not possible to calculate reliably structural energy differences associated with symmetry-lowering displacements of atoms. For example. ASA is sufficient for calculating bulk moduli (where the symmetry of the structure considered does not change during deformation) but it fails when calculating shear elastic constants or phonon frequencies. The reason is an inaccurate evaluation of the Coulomb and exchange-correlation energies with the spheroidized charge density within the ASA [21, 34]. As concluded by Tank et al. [35], this part of the total energy has the correct dependence on bond lengths, but not on bond angles. However, the non-spheroidized (full) charge density coming out of a conventional self-consistent LMTO-ASA calculations is quite accurate. Therefore the schemes calculating the Coulomb and exchange-correlation energy using full charge density [34, 36] are very successful in total energy evaluations. The third generation of the TB-LMTO method is being developed [15, 35] which should improve the calculations of total energy considerably.

The band structure energy, densities of states etc. obtained from the present LMTO-ASA calculations are also very similar to those obtained by full-potential calculations, even for low-symmetry configurations [21]. This suggests that the band-structure energy can reliably be calculated using the LMTO-ASA, which is relatively easy and fast, and the Coulomb and exchange-correlation energy may be represented by some semiempirical terms. As shown by Foulkes and Haydock [37], these semiempirical terms can be described, at least in principle, by pair potentials. This approach constitutes a basis for our semiempirical TB-LMTO-ASA approach.

6. SEMIEMPIRICAL TB-LMTO-ASA CALCULATIONS

For computer simulation of atomic configuration of extended defects in metals, semiempirical interatomic potentials have frequently been used (for a recent review, see e.g. [38]). These potentials usually do not include angular forces and directional bonds. Nowadays, bond-order potentials are being developed which include angular forces as well [39].

However, according to Sec. 5, we can use the TB-LMTO-ASA Hamiltonian as a basis for a semiempirical quantum-mechanical method of the TB type that includes angular forces. The total energy of the system may be written as [40]:

$$E_{tot} = E_{band}^{LMTO-ASA} + E_{AS}^{EMT} + E_{rep}. \tag{17}$$

Here $E_{band}^{LMTO-ASA}$ is the band-structure energy evaluated within the ASA, E_{AS}^{EMT} represents a correction to the ASA taken from the effective medium theory [41] and E_{rep} is a repulsive

term simulated by short-ranged pair potential.

If there are no adjustable parameters in the band term, the pair potential term may not be sufficient to describe variations of the total energy satisfactorily (e.g., we could not obtain the correct Cauchy pressures for elastic constants [42]). The presence of the E_{AS}^{EMT} makes it possible to remove this discrepancy.

The quantities E_{AS}^{EMT} and E_{rcp} contain adjustable parameters that are fitted to reproduce the equilibrium lattice constant, bulk modulus and shear moduli $C' = (C_{11} - C_{12})/2$ and C_{44} for the most stable structure of the unperturbed material. Angularly-dependent forces are included automatically.

Although the adjustable parameters in the terms E_{AS}^{EMT} and E_{rcp} are fitted to properties of the ground-state structure only (lattice constants, elastic moduli), the formula (17) yields correct description of the structural energy differences, energetics of large homogeneous deformations ($\sim$ 100 % and higher), phonon frequencies etc. [40, 43], i.e. also in more distant and non-equilibrium regions of the configurational space.
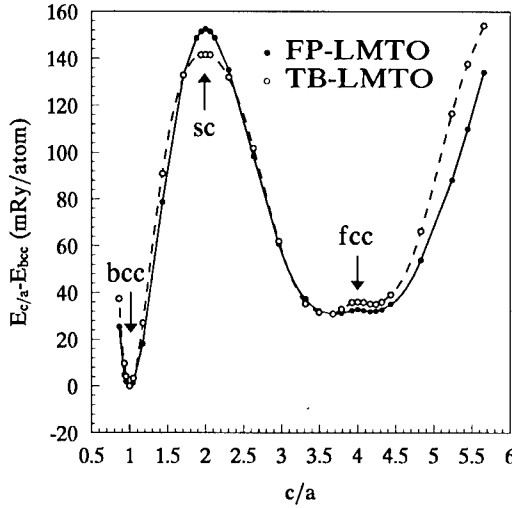
Fig. 6. The energy of trigonally deformed tungsten as a function of c/a ratio at constant volume.

Fig. 6 shows the total energy of trigonally deformed tungsten vs. c/a ratio calculated both by full-potential LMTO method and by the present semiempirical scheme. The c/a ratio is defined to be equal to 1 for bcc structure and 4 for fcc structure (at $c/a=2$, we obtain the simple cubic structure so that this deformation path connects continuously the bcc, simple cubic and fcc structures [44]). It may be seen from Fig. 6 that the full-potential LMTO curve is well reproduced by the present semiempirical approach including the features in the neighbourhood of the fcc structure. Comparisons of this type give us some confidence when using the present scheme for simulation of the structure of extended defects.

7. CONCLUSIONS

It was shown that the first-principles TB-LMTO-ASA method employing surface Green's function technique and/or coherent potential approximation is very useful for electronic structure calculations of complex systems and materials with defects. Provided the atomic configuration of the defective region is known, the method gives reliable electronic structure characteristics as for example local densities of states, charge densities etc. The LMTO-ASA total energy exhibits a correct dependence on bond lengths, but not on bond angles. Therefore, the first-principles TB-LMTO-ASA method cannot be used to calculate reliably structural energy differences associated with symmetry-lowering displacements of atoms and for relaxing defect configurations.

However, the TB-LMTO-ASA band structure energy can be combined with a semiempirical term describing the effect of electrons in the interstitial region and the repulsive term described by a pair potential. This approach reproduces very well the full-potential LMTO results for high-symmetry structures and their large deformation and, therefore, is very promising for atomistic simulations of extended defects.

REFERENCES

1. E. Lafon and C.C. Lin, Phys. Rev. **152**, 597 (1966).
2. W.A. Harrison, Electronic Structure and the Properties of Solids (Dover Publications, New York, 1989).
3. J. Callaway, Quantum Theory of the Solid State (Academic Press, Boston, 1991).
4. Solid State Physics, vol. **35**, edited by H. Ehrenreich, F. Seitz, and D. Turnbull, (Academic Press, New York, 1980).
5. A. Gonis, Green Functions for Ordered and Disordered Systems (North-Holland, Amsterdam, 1992).
6. R. Haydock and V. Heine, Comments Cond. Mat. Phys. **18**, 83 (1997).
7. R. Haydock, this Proceedings.
8. D.R. Bowler, M. Aoki, C.M. Goringe, A.P. Horsfield, and D.G. Pettifor, Modelling Simul. Mater. Sci. Eng. **5**, 199 (1997).
9. O.K. Andersen and O. Jepsen, Phys. Rev. Lett. **53**, 2571 (1984).
10. O.K. Andersen, O. Jepsen, and D. Glötzel, in Highlights of Condensed Matter Theory, edited by F. Bassani, F. Fumi, and M.P. Tosi (North-Holland, New York, 1985), p. 59.
11. O.K. Andersen, Z. Pawlowska, and O. Jepsen, Phys. Rev. B **34**, 5253 (1986).
12. O.K. Andersen, O. Jepsen, and M. Šob, in Electronic Band Structure and Its Applications, edited by M. Yussouff (Springer-Verlag, Berlin-Heidelberg, 1987), p. 1.
13. R. Zeller, P.H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, Phys. Rev. B **52**, 8807 (1995).
14. O.K. Andersen, O. Jepsen, and G. Krier, in Lectures on Methods of Electronic Structure Calculations, edited by V. Kumar, O.K. Andersen, and A. Mookerjee (World Scientific, Singapore, 1994), p. 63.

15. O.K. Andersen, G. Krier, R.W. Tank, C. Arcangeli, T. Dasgupta, and O. Jepsen, this Proceedings.
16. O.K. Andersen, Solid State Commun. **13**, 133 (1973).
17. O.K. Andersen, Phys. Rev. B **12**, 3060 (1975).
18. O.K. Andersen, in The Electronic Structure of Complex Systems, edited by P. Phariseau and W.M. Temmerman (Plenum, New York, 1984) p. 11.
19. H.L. Skriver, The LMTO Method (Springer, Berlin, 1984).
20. I. Turek, V. Drchal, J. Kudrnovský, M. Šob, and P. Weinberger, Electronic Structure of Disordered Alloys, Surfaces and Interfaces (Kluwer Academic Publishers, Boston-London-Dordrecht, 1997).
21. O.K. Andersen, A.V. Postnikov, and S.Yu. Savrasov, in Applications of Multiple Scattering Theory to Materials Science, edited by W.H. Butler, P.H. Dederichs, A. Gonis, and R.L. Weaver (Mater. Res. Soc. Symp. Proc. vol. **253**, Pittsburgh, PA, 1992), p. 37.
22. F. Garcia-Moliner and V.R. Velasco, Progr. Surf. Sci. **21**, 93 (1986); Theory of Single and Multiple Interfaces, (World Scientific, Singapore-New Jersey-London-Hong Kong, 1992).
23. B. Wenzien, J. Kudrnovský, V. Drchal, and M. Šob, J. Phys.: Condens. Matter **1**, 9893 (1989).
24. J. Kudrnovský, I. Turek, and V. Drchal, in Lectures on Methods of Electronic Structure Calculations, edited by V. Kumar, O.K. Andersen, and A. Mookerjee (World Scientific, Singapore, 1994), p. 231.
25. J. Kudrnovský, I. Turek, V. Drchal, and M. Šob, in Stability of Materials, edited by A. Gonis, P.E.A. Turchi, and J. Kudrnovský (Plenum Press, New York, 1996), p. 237.
26. G. Ackland and R. Thetford, Phil. Mag. A **56**, 15 (1987).
27. S. Crampin, D.D. Vvedensky, J.M. MacLaren, and M.E. Eberhart, Phys. Rev. B **40**, 3413 (1989).
28. A.J. Freeman and R. Wu, J. Magn. Magn. Mater. **100**, 497 (1991).
29. M. Weinert and S. Blügel, in Magnetic Multilayers, edited by L.H. Bennett and R.E. Watson (World Scientific, Singapore, 1994), p. 51.
30. G.A. Mulhollan, R.L. Fink, and J.L. Erskine, Phys. Rev. **44**, 2393 (1991).
31. C. Liu and S.D. Bader, Phys. Rev. B **44**, 12062 (1991).
32. I. Turek, J. Kudrnovský, M. Šob, V. Drchal, and P. Weinberger, Phys. Rev. Lett. **74**, 2551 (1995).
33. S. Blügel, Phys. Rev. B **51**, 2025 (1995).
34. O.K. Andersen, M. Methfessel, C.O. Rodriguez, P. Blöchl, and H.M. Polatoglou, in Atomistic Simulations of Materials: Beyond Pair Potentials, edited by V. Vitek and D.J. Srolovitz (Plenum, New York-London, 1989), p. 1.
35. R.W. Tank, C. Arcangeli, G. Krier, O.K. Andersen, and O. Jepsen, in Properties of Complex Inorganic Solids, edited by A. Gonis, A. Meike, and P.E.A. Turchi (Plenum, New York-London, 1997), p. 233.
36. L. Vitos, J. Kollár, and H.L. Skriver, Phys. Rev. B **55**, 13521 (1997).
37. W.M.C. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989).
38. V. Vitek, in Stability of Materials, edited by A. Gonis, P.E.A. Turchi, and J. Kudrnovský (Plenum, New York, 1996), p. 53.
39. A.P. Horsfield, A.M. Bratkovsky, M. Fearn, D.G. Pettifor, and M. Aoki, Phys. Rev. B **53**, 12964 (1996).
40. M. Šob, V. Vitek, and Y. Oh, in Computational Methods in Materials Science, edited by J.E. Mark, M.E. Glicksman, and S.P. Marsh (Mater. Res. Soc. Symp. Proc. vol. **278**, Pittsburgh, PA, 1992), p. 205.
41. K.W. Jacobsen, J.K. Nørskov, and M.J. Puska, Phys. Rev. B **35**, 7423 (1987).
42. M. Šob and V. Vitek, in Stability of Materials, edited by A. Gonis, P.E.A. Turchi, and J. Kudrnovský (Plenum, New York, 1996), p. 449.
43. M. Šob and V. Vitek, to be published.
44. M. Šob, L.G. Wang, and V. Vitek, Comput. Mat. Sci. **8**, 100 (1997).

An ab initio two-center tight-binding approach to simulations of complex materials properties

Th. Frauenheim, D. Porezag, M. Elstner, G. Jungnickel, J. Elsner, M. Haugk, A. Sieck
and G. Seifert*, *Technische Universität, Institut für Physik, D-09107 Chemnitz, Germany,*
* *Technische Universität, Institut für Theoretische Physik, D-01069 Dresden, Germany*

ABSTRACT

We describe the *ab initio* construction of two-center tight-binding (TB) hamiltonians, which at a properly chosen input density upon non-selfconsistent solution of the related Kohn-Sham equations transform the energy within density-functional theory (DFT) into a tight-binding-like expression. In cases, where the electron density of the interacting many-atom structure in good approximation may be represented as a sum of atomic-like densities, the method has been shown to operate highly transferable, being particularly successful in determining the properties of low-energy silicon clusters, in predicting the structure and vibrational signatures of fullerene oligomers, amorphous carbons and carbon nitrides and in simulating elementary growth reactions on diamond surfaces. The uncertainties within the standard non-SCF DF-TB-variant, however, increase if the chemical bonding is controlled by a delicate charge balance between different atomic constituents, as e.g. in organic molecules and in polar semiconductors. Therefore, we extend the standard TB-approach to the operation in a selfconsistent-charge mode *(SCC-DFTB)* in order to improve total energies, forces, and transferability in the presence of considerable long-range Coulomb interactions. By using a variational technique, we derive a transparent and readily calculable expression for the iterative modification of Hamiltonian matrix elements and show, that the final energy is a second order approximation to the total energy in density-functional theory, *see M. Elstner et al., this Symposium.* First successful applications to surface studies of GaAs and dislocation modeling in GaN will be presented.

I. INTRODUCTION

It has been shown that the TB-approach in general may be understood as a stationary approximation to density-functional theory (DFT) [1, 2, 3, 4, 5]. Central features of the common methodology, namely, non-selfconsistent treatment of the *Kohn-Sham* equations and the exploitation of pairwise repulsive interactions are strongly related to an appropriate "educated guess" for the inital charge density of the system. This makes the total energy readily computable and second-order effects arising from a charge redistribution negligibly small. Further simplifications that frequently enter calculations are related to the introduction of semi-empirical parameters or functionals into the total energy expression [6, 7]. The definition of the Hamiltonian matrix elements based on adjustable values has been applied very successfully to many problems. However, such methods often fail in bonding situations that are not covered by the parametrization space. These are exactly the cases of real interest.

In order to completely avoid the difficult parametrization within a multiconfigurational space, more sophisticated, yet efficient, TB-schemes have recently been developed. These methods include the TB-LMTO (linear-muffin-tin-orbitals) method [8], the Hartree-Fock-based TB [9], a successful DFT-parametrization of TB [10], the *ab initio* multicenter TB [5], and our DF-based *(two-center)* TB approach [11]. Here, the Hamiltonian matrix elements are explicitly calculated within a non-orthogonal basis of atomic orbitals. These schemes yield accurate results for a broad range of bonding situations, for which the superposition of overlapping atom-like densities serves as a good approximation for the many-atom structure. Nevertheless, problems naturally arise if a delicate charge balance is required for establishing chemical bonding between different types of atoms. In such cases, an adjustment of the charge distribution via a selfconsistent (SCF) procedure may take place. Hence, there is a need to extend the TB formalism in order to improve its transferability without discarding the simplicity, speed and efficiency which make it so useful.

Within the framework of standard empirical TB theory, several proposals have been made to generalize the TB-energies by explicitly considering interatomic electron interactions within a standard Hubbard approach [12, 13, 14, 15]. Skriver and Rosengaard use an efficient selfconsistent Green's-function technique based on the LMTO-method within the tight-binding and atomic-sphere approximation [16]. Tsai et al. [17] improved the *ab initio* multicenter TB scheme [5] in order to account for charge transfer. By using the Ewald technique, they added and subtracted to each atom a Gaussian charge distribution and solved the Schrödinger equation iteratively, to determine the selfconsistent atomic charges. Demkov et al. [18] used a different approach to modify this TB scheme by including long ranged *Hartree* contributions, which leads to a selfconsistent adjustment of site-dependent occupation numbers.

Here, we focus on a systematic extension of the tight-binding formalism and of our DF-TB scheme [3, 11] in order to derive a generalized selfconsistent charge (SCC) methodology. This differs from previous approaches since we base the modification of the TB total energy expression on a second-order expansion of the *Kohn-Sham* energy functional [19] with respect to density fluctuations. The new methodology ensures a proper distribution of the charge and overcomes the requirement of local charge neutrality [4], especially in multi-component systems. In maintaining the simple two-center picture, the new scheme can be easily incorporated into any standard TB-method. We demonstrate the improvements considering properties of molecular systems, where the non-SCC scheme failed, and point out recent successful applications to polar semiconductors.

II. DENSITY-FUNCTIONAL BASED TIGHT-BINDING, DF-TB

The total energy of a system of M electrons in the field of N nuclei at positions $\vec{R}$ may be written within DFT as a functional of a charge density $n(\vec{r})$:

$$E = \sum_i^{occ} \langle \Psi_i | -\frac{\Delta}{2} + V_{ext} + \frac{1}{2}\int' \frac{n(\vec{r}')}{|\vec{r}-\vec{r}'|} | \Psi_i \rangle + E_{xc}[n(\vec{r})] + \frac{1}{2}\sum_{\alpha,\beta}^{N} \frac{Z_\alpha Z_\beta}{|\vec{R}_\alpha - \vec{R}_\beta|} ,$$

(1)

where the first sum is over occupied *Kohn-Sham* eigenstates Ψ_i, the second term is the exchange-correlation (XC) contribution, and the last term covers the ion-ion core repulsion, E_{ii}. We now rewrite the total energy in order to transform the leading matrix elements. We first substitute the charge density in (1) by a superposition of a reference or input density $n_0' = n_0(\vec{r}')$ and a small fluctuation $\delta n' = \delta n(\vec{r}')$, $\int d\vec{r}'$ is expressed by $\int'$:

$$\begin{aligned} E &= \sum_i^{occ} \langle \Psi_i | -\frac{\Delta}{2} + V_{ext} + \int' \frac{n_0'}{|\vec{r}-\vec{r}'|} + V_{xc}[n_0] | \Psi_i \rangle - \frac{1}{2}\iint' \frac{n_0'(n_0 + \delta n)}{|\vec{r}-\vec{r}'|} \\ &\quad - \int V_{xc}[n_0](n_0 + \delta n) + \frac{1}{2}\iint' \frac{\delta n'(n_0 + \delta n)}{|\vec{r}-\vec{r}'|} + E_{xc}[n_0 + \delta n] + E_{ii} . \end{aligned}$$

(2)

The second term in this equation corrects for the double-counting of the new *Hartree*, the third term for the new XC contribution in the leading matrix element and the fourth term comes from dividing the full *Hartree* energy in (1) into a part related to n_0 and to δn.

Finally, we expand E_{xc} at the reference density and obtain the total energy correct to second order in the density fluctuations. Note that the terms linear in δn cancel each other at any arbitrary input density n_0:

$$\begin{aligned} E &= \sum_i^{occ} \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle - \frac{1}{2}\iint' \frac{n_0' n_0}{|\vec{r}-\vec{r}'|} + E_{xc}[n_0] - \int V_{xc}[n_0]n_0 + E_{ii} \\ &\quad + \frac{1}{2}\iint' \left(\frac{1}{|\vec{r}-\vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta n \, \delta n'} \Big|_{n_0} \right) \delta n \, \delta n'. \end{aligned}$$

(3)

III. ZERO-th ORDER NON-SCC APPROACH

The traditional non-SCC TB-approach is to preserve the first 5 terms in this final equation, with $\hat{H}_0$ as the Hamiltonian operator resulting from an input density n_0. As usual, a frozen-core approximation is applied to reduce the computational efforts by only considering the valence orbitals. The *Kohn-Sham* equations are then solved non-selfconsistently and the second-order correction is neglected. The contributions in (3) that depend on the input density n_0 only and the core-core repulsion are taken to be a sum of one- and two-body potentials [2]. The latter, denoted by E_{rep}, are strictly pairwise, repulsive and short-ranged. The total energy then reads

$$E_0^{TB} = \sum_i^{occ} \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle + E_{rep}. \tag{4}$$

To solve the *Kohn-Sham* equations, the single-particle wave functions Ψ_i within an LCAO-*ansatz* are expanded into a suitable set of localized atomic orbitals φ_ν,

$$\Psi_i(\vec{r}) = \sum_\nu c_{\nu i} \varphi_\nu(\vec{r} - \vec{R}_k) , \tag{5}$$

As described earlier [11], we employ confined Slater-type atomic orbitals. These are determined by solving a modified Schrödinger equation for a free atom within SCF-LDA calculations. As an approximation, we write the one-electron potential of the many-atom structure as a sum of spherical atomic contributions:

$$V_{eff}(\vec{r}) = V_{ext} + \int' \frac{n_0'}{|\vec{r} - \vec{r}'|} + V_{xc}[n_0] = \sum_k V_0^k(|\vec{r} - \vec{R}_k|) , \tag{6}$$

where V_0 is the Kohn–Sham potential of a neutral pseudoatom due to its *confined* electron density. By applying the variational principle to the zero-th order energy functional (4), we obtain the *Kohn-Sham* equations which, finally, within the pseudoatomic basis, transform into a set of algebraic equations.

$$\sum_\nu^M c_{\nu i}(H_{\mu\nu}^0 - \varepsilon_i S_{\mu\nu}) = 0, \quad \forall \mu, i, \tag{7}$$

$$H_{\mu\nu}^0 = \langle \varphi_\mu | \hat{H}_0 | \varphi_\nu \rangle; \quad S_{\mu\nu} = \langle \varphi_\mu | \varphi_\nu \rangle ; \quad \forall \mu \epsilon \alpha, \ \nu \epsilon \beta. \tag{8}$$

Consistent with (6) we neglect several contributions to the Hamiltonian matrix elements $H_{\mu\nu}$ [3] yielding:

$$H_{\mu\nu}^0 = \begin{cases} \varepsilon_\mu^{\text{free atom}} & \text{if } \mu = \nu \\ \langle \varphi_\mu^A | \hat{T} + V_0^A + V_0^B | \varphi_\nu^B \rangle & \text{if } A \neq B \\ 0 & \text{otherwise} \end{cases} . \tag{9}$$

Since indices A and B indicate the atoms on which the wavefunctions and potentials are centered, only two-center Hamiltonian matrix elements are treated and explicitly evaluated in combination with the two-center overlap matrix elements. As follows from (9), the eigenvalues of the free atom serve as diagonal elements of the Hamiltonian, thus guaranteeing the correct limit for isolated atoms.

By solving the general eigenvalue problem (7), the first term in (4) becomes a simple summation over all occupied Kohn-Sham orbitals ε_i (occupation number n_i), while E_{rep} can easily be determined as a function of distance by taking the difference of the SCF-LDA cohesive and the corresponding TB band structure energy for a suitable reference system:

$$E_{rep}(R) = E_{LDA}^{SCF}(R) - \sum_i^{occ} n_i \varepsilon_i(R) . \tag{10}$$

93

Interatomic quantum forces for molecular-dynamics applications can easily be derived from an explicit calculation of the gradients of the total energy at the considered atom sites:

$$M_k \ddot{\vec{R}}_k = -\frac{\partial E_0^{TB}}{\partial \vec{R}_k} = -\sum_i^{occ} \left[\frac{\partial n_i}{\partial \vec{R}_k} \varepsilon_i + n_i \frac{\partial \varepsilon_i}{\partial \vec{R}_k} \right] - \sum_{k' \neq k} \frac{\partial E_{rep}(|\vec{R}_k - \vec{R}_{k'}|)}{\partial \vec{R}_k}, \qquad (11)$$

where the electronic part of the forces reads

$$\frac{\partial \varepsilon_i}{\partial \vec{R}_l} = \sum_\mu \sum_\nu c_{\mu i} c_{\nu i} \left[\frac{\partial H_{\mu\nu}^0}{\partial \vec{R}_l} - \varepsilon_i \frac{\partial S_{\mu\nu}}{\partial \vec{R}_l} \right]. \qquad (12)$$

IV. RESULTS AND TRANSFERABILITY

The just dicussed standard DF-TB-approach has been successfully applied to various problems in different systems and materials, covering carbon, silicon & germanium modifications, boron and carbon nitrides, silicon carbide and oxide and even III-V semiconductors. Provided a clever guess of the initial or input charge density of the system is found, the energies and forces are correct in second order of occuring charge density fluctuations. Furthermore, the once determined short-range two-particle repulsion (using a proper reference system) can be shown to operate transferable in very different bonding situations considering various scale systems. In the following we will demonstrate this with a few examples, finally approaching the limits of non-selfconsistency.

Prediction of low-energy silicon clusters /18/

We have used the non-SCC DF-TB-method to determine new lowest-energy structures for silicon clusters with 9 to 14 atoms. The confined Slater-type atomic orbitals were obtained by solving the Kohn-Sham equation for a single silicon atom with an additional potential $V(r) = (r/r_0)^2$, with $r_0 = 3.8 a_B$. We restrict ourself to a minimum s,p-basis set to allow fast molecular-dynamic simulations.

The tetrahedral Si_5 cluster was choosen to determine the shortrange repulsive potential in the tight-binding total energy expression. Although this structure is not the ground-state of Si_5 it represents a locally stable energy minimum. We set the repulsive potential equal to zero for $r \geq r_c = 5.2 a_B$.

For small silicon clusters with up to 8 atoms several highly accurate *ab initio* calculations has been performed, which, together with experimental IR and Raman spectra, have established the geometries and electronic states for clusters in this size range [21]. For Si_2 to Si_8 we obtain within the DF-TB-method the same equilibrium structures as Fornier *et al.* [22] and Pederson *et al.* [23]. Overall, there is good agreement between the SCF-LDA and DF-TB methodologies for geometrical parameters and binding energies for these small clusters. The deviations in bond lengths and angles are smaller than 10%. The deviations in cohesive energies are smaller than 4% for all clusters larger than Si_2.

Considering larger clusters with 9 to 14 atoms, we have started our search for the equilibrium structures with different seed clusters obtained by edge or face capping of smaller stable clusters or with clusters taken from the literature. We have optimized these structures with our DF-TB method by applying either a stochastic molecular-dynamic quenching or conjugate gradient relaxation until the maximum force on every atom dropped below 10^{-4} Hartree/Bohr. At least the two most stable DF-TB structures were then relaxed with the SCF-LDA code of Pederson *et al.* [24] using a 6s5p3d basis set.

Since the DF-TB method does not take spin into account, we have shifted the DF-TB energies by the spin-polarization energy of 0.656 eV. The cohesive energies and highest occupied to lowest unoccupied molecular orbital (HOMO-LUMO) gaps for the lowest energy clusters are summarized in Table I. Si_7, Si_{10} and Si_{14} are more stable than their neighbors. The order of the DF-TB determined cohesive energies of the lowest energy structures for one cluster size agrees with the

Cluster	(sym.)	E_{TB}	E_{scf}	Gap_{TB}	Gap_{scf}
Si2	$(D_{\infty h})$	-1.936	-1.780	0.000	0.000
Si3	(C_{2v})	-2.983	-2.965	1.874	1.008
Si4	(D_{2h})	-3.488	-3.541	1.453	1.075
Si5	(D_{3h})	-3.766	-3.825	1.702	1.976
Si6	(C_{2v})	-3.925	-4.041	1.337	2.106
Si7	(D_{5h})	-4.063	-4.187	1.511	2.097
Si8a	(C_{2h})	-4.071	-4.122	1.155	1.419
Si8b	(C_{2v})	-4.087	-4.072	0.843	1.086
Si9a	(C_{2v})	-4.176	-4.234	1.904	1.988
Si9b	(C_{2v})	-4.126	-4.183	1.846	1.551
Si9c	(D_{3h})	-4.031	-4.097	0.173	0.397
Si10a	(C_{3v})	-4.242	-4.357	1.706	2.125
Si10b	(T_d)	-4.129	-4.286	3.405	2.136
Si11a	(C_s)	-4.203	-4.274	1.214	1.041
Si11b	(E)	-4.210	-4.262	1.009	0.922
Si11c	(C_s')	-4.205	-4.259	1.330	1.073
Si12a	(C_s)	-4.228	-4.274	0.925	0.593
Si12b	(E)	-4.250	-4.267	0.862	0.940
Si13a	(C_{3v})	-4.204	-4.305	1.451	1.606
Si13b	(C_{2v})	-4.277	-4.291	1.332	0.787
Si14a	(C_s)	-4.328	-4.372	1.531	1.774
Si14b	(C_s)	-4.300	-4.332	1.029	1.323
Si14c	(D_{3h})	-4.283	-4.253	1.095	0.896
Si14d	(C_{3v})	-4.179	———	0.000	——

Table I: Binding energies with respect to spin polarized atoms in eV/atom and HOMO/LUMO gap in eV for silicon clusters as calculated within DF-TB and scf-LDA.

SCF-LDA results for all smaller clusters and Si_9, Si_{10} and Si_{14} but is reversed for Si_8, Si_{11}, Si_{12} and Si_{13}. Please note that for the three latter clusters the differences in the SCF cohesive energies are smaller than 0.02 eV/atom, which certainly is below the accuracy of the DF-TB approach. The variation of the HOMO-LUMO gap as calculated within DF-TB with increasing cluster size, n, is in good agreement with the SCF-LDA calculations for n > 7. The quantitative differences for Si_8 to Si_{14} are at maximum of 20%, whereas the deviations increase with decreasing cluster size. Note that an accurate description of unoccupied orbitals often fails due to the use of a minimum basis set. We obtained the following results for Si_9, Si_{11}, Si_{12}, Si_{13} and Si_{14}:

Si_9: We found two stacked distorted rhombi with an additional atom capped on top (Si_{9a}) to be the lowest-energy cluster. This structure has C_{2v} symmetry and is 0.05 eV/atom more stable than the distorted tricapped trigonal prism Si_{9b}, first proposed by Ordejon et al. [25] as the most stable structure of Si_9.

Si_{11}: We confirm a structure, Si_{11a}, proposed by Lee et al. [26] using a TB-method to be the most stable Si_{11} cluster. This structure may be seen as a distorted pentacapped trigonal prism. The relaxation out of different starting structures for Si_{11} with the TB-method has also spawned another cluster, Si_{11b}, having no symmetry at all. This cluster is only 0.01 eV/atom less stable than Si_{11a}. It is a distorted 4-5 prism, with two atoms capping the fivefold ring.

Si_{12}: Recently Ramakrishna et al. have presented results of their extensive search for the ground state of Si_{12} [27]. They report on six isomers, which differ by only 0.02 eV/atom in cohesive energies, as calculated within the DF-LDA. These isomers can all be described either as hexacapped trigonal prisms or antiprisms with different faces capped. Considering their isomer 2 (Si_{12a}), a hexacapped distorted trigonal prism with (C_s) symmetry, we obtain a cohesive energy of -4.274 eV/atom. We also find another, nonsymmetric, structure Si_{12b} to be only 0.01 eV/atom less stable than their isomer 2.

Si_{13}: The most stable structure of Si_{13} consists of four stacked triangles with a cap on top. Recent QMC [28] and DF-LDA [23, 29] calculations found this structure Si_{13a} to be favoured

against the icosahedral form of Si_{13}. From the present search we predict the cohesive energy of another structure to be very close to that of the lowest energy cluster. This structure, Si_{13b}, has C_{2v} symmetry and can be described as a distorted tricapped trigonal prism with an additional rhombus capped on one edge of the prism. The energy difference between these two structures is only 0.014 eV/atom.

Si_{14}: Only a few structures for Si_{14} have been presented in the literature so far. We suggest a stacking sequence of two distorted rhombi, one fivefold ring and an atom on top (Si_{14a}) as a candidate for the equilibrium structure. Another local stable isomer is the octacapped trigonal prism with D_{3h} symmetry (Si_{14c}). This geometry is 0.12 eV/atom less stable than the lowest energy candidate.

To aid in the experimental identification of these clusters, we have computed the full vibrational spectra of the structures, along with the Raman activities, IR intensities and static polarizabilities, using SCF-DF theory within the LDA. A detailed description of the results can be found in [20].

Energetic stability and structure of amorphous carbon nitride /29/

The behaviour of nitrogen in carbon systems is an area of intense interest within the materials science research community. At high N incorporation rates, the theoretical prediction of the possible formation of a superhard C_3N_4-phase [32], has spawned an, as yet, unsuccessful worldwide effort to fabricate it. Various deposition techniques have produced amorphous films with a nitrogen content below 40%, which is found to be an upper limit for nitrogen incorporation. Infrared (IR) and Raman spectra show that in films with nitrogen content over 30%, in addition to single and double bonds some triple bonding is also present [33]. Simultaneously the mass density of 1.8 to 2.3 g/cm^3 is usually found to be much lower than the desired 3.5 g/cm^3 in the low-compressibility pases [34].

In reasoning these problems on an atomistic level, we have investigated the chemistry of nitrogen incorporation into carbon systems by means of the DF-TB molecular-dynamics (MD) scheme. Using MD annealing techniques, we generated amorphous carbon nitride structures, a-CN, with various stoichiometries (0 to 57 %) and densities (1.5 to 4.0 g/cm^3).

Focusing on high densities, we observe clear trends counteracting the formation of a low-compressibilty phase: (i) N-incorporation strongly catalyzes C-undercoordination, which in turn (ii) causes the nitrogens to develop paracyanogen-like (CN-double and -triple) bonding. & (iii) The most favorable densities for a-CN are far lower than that desired by hard crystalline phases. A closer analysis of the high density models with 14% N reveals the nitrogens to be fairly randomly distributed over the cell, bonding to both sp^3- and sp^2-like carbon. Most of the N's have three C-neighbors and form single bonds. Approximately one half of them are sp^3-hybridized with one filled lone-pair orbital, while the other half shows sp^2-hybridization with C–N–C bond angles close to 120°. These models can be regarded as heavily N-doped a-C, in which the nitrogens are merely substituted into the C-network.

The microstructure of the models changes considerably, however, at higher N-incorporation rates. The structural properties are governed to an ever increasing extent by the C–N-bonding, which is highly complex, since both atoms can exhibit several hybridizations, depending on the local atomic environment. The changes can be best understood by considering the nitrogens, which should be preferably sp^2-hybridized as in the β-phase with three single-bonded C-neighbors. With increasing N-content, we observe a pronounced reduction of the N–C coordination number, decreasing from almost 3 at low N-concentration to approximately 2 at 57% N. The overall C-coordination simultaneously decreases from 3.91 to 3.61 at 3.5 g/cm^3. The related lowering in the carbon sp^3-fraction versus N-content is plotted in Fig.1 for the three highest densities.

In the nitrogen-free case, the sp^3-content rises from approximately 65% at 3.0 g/cm^3 to still less than 100 % (90 & 94 %) at 3.5 and 4.0 g/cm^3. The undercoordinated atoms preferentially arrange in pairs, lowering the total energy by forming additional π-clusters. This agrees with former studies using a slightly different annealing scheme [35] and with predictions of non-SCF ab $initio$ MD[36]. The addition of nitrogen lowers the sp^3-carbon content drastically. This tendency
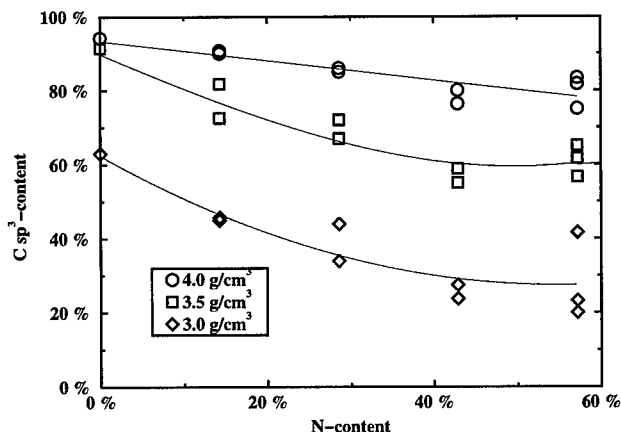
Figure 1: Carbon sp^3-fraction versus N-content. *(lines are just drawn to guide the eye.)*

holds true up to a nitrogen concentration of 43%, where we find considerably decreased fractions of 26%, 57% and 78% sp³-hybridized carbon at ρ=3.0, 3.5 and 4.0 g/cm³, respectively. At the C_3N_4-stoichiometry, the values are essentially unaltered. By comparing the models with low and high N-fractions at crystalline density of 3.5 g/cm³, the difference is immediately obvious: while the former exhibits a compact, three dimensional cross-linked network, the connectivity in the nitrogen-rich cell is considerably reduced; there are even some chain-like and layered structure elements. Such trends will strongly counteract the successful synthesis of low-compressibility C_3N_4. At the latter stoichiometry we find mainly three types of bonding: 1) 15-20% of the nitrogens have three singly-bonded neighbors, again with approximately one half in the most favourable sp²-hybridization; the remaining are sp³-like with a doubly occupied lone pair orbital. 2) 35-45% are bonded to only two carbons, most to one sp³- and one sp²-C, forming one double and one single bond. Finally, 3) 25-35% have only one neighboring C-atom, forming short cyanide *(triple-)* bonds. The type 2) and 3) bonding configurations are well-known from paracyanogen, a polymerized form of cyanogen, which exhibits conjugated C–N bonds and cyanide groups attached to the network. The occurance of paracyanogen-like bonding at even higher densities can explain why the IR and Raman spectra of films with an N-content above 30% show not only peaks from single and double CN-bonds, but also a peak at 2150 cm⁻¹, which is attributable to cyanide bonds. A physical intepretation of the mixed nature of the C–N bonding in a-C_3N_4 can be gained by noting that bonding configurations 1), 2) & 3) are almost isoenergetic. We therefore arrive at a surprisingly simple and powerful argument explaining the difficulty in fabricating c-C_3N_4: a pure crystalline low-compressibility form is unlikely to occur when the bonding configuration is in competition with the two other, equally viable bonding configurations during the growth process.

IV. THE SELFCONSISTENT-CHARGE EXTENSION, SCC-DFTB

All the so far discussed applications may be seen as examples, where the electron density of the interacting many-atom structures in good approximation may be represented as a sum of atomic-like densities. The uncertainties within the standard DF-TB-variant, however, increase if the chemical bonding is controlled by a delicate charge balance between different atomic constituents, as e.g. in organic molecules and in polar semiconductors. Therefore, we have extended the

standard TB-approach to the operation in a selfconsistent-charge mode *(SCC-DFTB)* in order to improve total energies, forces, and transferability in the presence of considerable long-range Coulomb interactions. We are starting from equation 3 and now explicitely consider the second order term in the density fluctuations. We first decompose $\delta n(\vec{r})$ into atom-centered contributions which we expand in a series of radial and angular functions. Truncating the multipole expansion after the monopole term accounts for the most important contributions of the charge transfer yielding an approximate expression for the second order term, for a detailed discussion see M. Elstner et al. [37] and J. Elsner et al. [38].

Finally, the DFT total energy (3) is conveniently transformed into a transparent TB-form,

$$E = \sum_i^{occ} \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle + \frac{1}{2} \sum_{\alpha,\beta}^N \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta + E_{rep}, \tag{13}$$

where $\gamma_{\alpha\beta} = \gamma_{\alpha\beta}(U_\alpha, U_\beta, |\vec{R}_\alpha - \vec{R}_\beta|)$ is analytically evaluated as the exact Coulomb interaction of two exponentially decaying charge contributions that integrate to the correct atomic net charges and reproduce the Hubbard-parameters U_α of the isolated atoms in the limit $|\vec{R}_\alpha - \vec{R}_\beta| = 0$, calculated at the LDA-level. This approach is consistent with the use of Slater-type orbitals as basis and guarantees a fast transition to Coulomb-like behavior for increasing distances, see also [38, 39]. As discussed earlier, the contribution due to $\hat{H}_0$ depends only on n_0 and is therefore exactly the same as in the previous non-SCC studies [11]. However, since the atomic charges depend on the one-particle wave functions Ψ_i, now a selfconsistent procedure is required to find the minimum of expression (13).

To solve the *Kohn-Sham* equations, the single-particle wave functions Ψ_i again are expanded into a suitable set of localized atomic orbitals φ_ν, compare equation (5). In accord with the previous scheme [11], we employ confined Slater-type atomic orbitals. These are determined by solving a modified Schrödinger equation for a free atom within SCF-LDA calculations. By applying the variational principle to the energy functional (13), we obtain the *Kohn-Sham* equations which, within the pseudoatomic basis, transform into a set of algebraic equations. We employ the Mulliken charge analysis for estimating the charge fluctuations $\Delta q_\alpha = q_\alpha - q_\alpha^0$,

$$q_\alpha = \frac{1}{2} \sum_i^{occ} n_i \sum_{\mu \in \alpha}^N \sum_\nu \left(c_{\mu i}^* c_{\nu i} S_{\mu\nu} + c_{\nu i}^* c_{\mu i} S_{\nu\mu} \right) \tag{14}$$

and obtain:

$$\sum_\nu^M c_{\nu i}(H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) = 0, \quad \forall \mu, i, \tag{15}$$

$$H_{\mu\nu} = \langle \varphi_\mu | \hat{H}_0 | \varphi_\nu \rangle + \frac{1}{2} S_{\mu\nu} \sum_\xi^N (\gamma_{\alpha\xi} + \gamma_{\beta\xi}) \Delta q_\xi \tag{16}$$

$$= H_{\mu\nu}^0 + H_{\mu\nu}^1 \; ; \quad S_{\mu\nu} = \langle \varphi_\mu | \varphi_\nu \rangle \; ; \quad \forall \mu \epsilon \alpha, \; \nu \epsilon \beta.$$

Since the overlap matrix elements $S_{\mu\nu}$ generally extend over a few nearest neighbor distances, they introduce multiparticle interactions. The second-order correction due to charge fluctuations is now represented by the Mulliken charge dependent contribution $H_{\mu\nu}^1$ to the matrix elements $H_{\mu\nu}$.

As in previous studies [11], we determine the short-range repulsive pair potential E_{rep} as a function of distance by taking the difference of the SCF-LDA cohesive energy and the corresponding TB band structure energy (which is now modified by the Coulomb correction) for a suitable reference structure. Since charge transfer effects are now considered explicitely, the transferability of E_{rep} is significantly improved compared to the non-SCC approach. Again, we find a simple analytic expression for the interatomic forces by taking the derivative of the final SCC-DFTB energy (13) with respect to the nuclear coordinates,

$$\vec{F}_k = -\sum_i n_i \sum_{\mu\nu} c_{\mu i} c_{\nu i} \left(\frac{\partial H_{\mu\nu}^0}{\partial \vec{R}_k} - \left(\varepsilon_i - \frac{H_{\mu\nu}^1}{S_{\mu\nu}} \right) \frac{\partial S_{\mu\nu}}{\partial \vec{R}_k} \right) - \Delta q_\alpha \sum_\xi^N \frac{\partial \gamma_{\alpha\xi}}{\partial \vec{R}_k} \Delta q_\xi - \frac{\partial E_{rep}}{\partial \vec{R}_k}. \tag{17}$$

V. IMPROVED TRANSFERABILITY, RESULTS

In the remainder of this paper, we shall concentrate on presenting results of first successful applications of the SCC-DFTB scheme to a wide range of systems which are of interest in chemistry, biology, and physics. In particular, we demonstrate the improvements of the method as compared to conventional DF-TB and *ab initio* calculations.

Organic and biological systems /37/

For our first benchmark, we have calculated the reaction energies of 36 processes between small closed shell molecules containing oxygen, nitrogen, carbon, and hydrogen from [40]. We have found a mean absolute deviation from experiment of 12.5 kcal/mol for the SCC-TB, compared to 11.1 kcal/mol for the DFT-LSD calculations. Further, considering the optimized geometries of a 63 organic molecules test set from Ref. [41], the mean absolute deviations from experiment in the bond lenghts and bond angles are $\Delta R = 0.010 \mathring{A}$ and $\Delta \theta = 1.95°$ [39] respectively, compared to $\Delta R = 0.017 \mathring{A}$ and $\Delta \theta = 2.01°$ by using the semi-empirical AM1 Method [41]. The improvement over the non-SCC treatment is impressively demonstrated for systems with a delicate counterbalance between ionic and covalent bonding contributions, as e.g. in formic acid (see M. Elstner et al., this volume). The DF-TB method overestimates the equalization of single and double bonds in the amide and carboxyl groups. This is exclusively due to too much charge flow (of nearly one electron) from carbon to oxygen, clearly indicating the need for a selfconsistent charge redistribution. SCC can considerably improve vibrational frequencies of very simple molecules like CO_2, in which charge transfer dominates. The symmetric and antisymmetric stretching modes ($\sum_g$ & $\sum_u$) change from 1458 & 1849 cm^{-1} in DF-TB to 1348 & 2305 cm^{-1} in SCC-DFTB (which is in good agreement with the experimental values, 1333 & 2349 cm^{-1}). Frequencies have further been tested for a series of 33 standard organic molecules yielding 6.4 % mean absolute deviation of vibrational frequencies from the experiment [39]. These are very promising results for future applications of the SCC-DFTB scheme to investigations of the geometry, dynamical behavior electronic structure of large biomolecules [43].

GaAs surface reconstruction /42/

The GaAs $(\overline{111})$ $\sqrt{19} \times \sqrt{19}$ surface periodicity can be observed by heating [45, 46] or annealing [47] the sample at $\approx 500°C$. A significant desorption of surface As is reported at the transition from 2×2 to $\sqrt{19} \times \sqrt{19}$ periodicity. Moreover, Woolf *et al.* [48] found that the $\sqrt{19} \times \sqrt{19}$ periodicity does not exist under a strong As$_4$ flux, clearly showing that this particular surface should exist in a Ga rich environment.

Several structural models have already been suggested to explain the $\sqrt{19} \times \sqrt{19}$ periodicity [47, 49]. However, a total energy calculation by standard SCF–LDA schemes is still lacking due to the computational expenses. Here we are discussing the first conclusive calculations for the formation energies of various models for the GaAs $(\overline{111})$ $\sqrt{19} \times \sqrt{19}$ reconstruction together with previously examined models for 2×2 reconstructions [47, 49].

The relative stabilities of two structures having different numbers of Ga and As atoms depend on the reservoir with which the atoms are exchanged in the structural transition. Questions of thermodynamic stability are therefore posed within the context of the atomic chemical potentials [50]. The surfaces were modeled by two dimensional slabs. To obtain absolute surface energies we use an energy density formalism which is described in detail in Ref. [51].

Kaxiras *et al.* [49] proposed a structural model consisting of Ga trimers and three–fold coordinated As atoms at the surface. In this structure (see Ref.[49] for a figure) all As lone pairs are filled and $\frac{3}{4}$ of an electron would have to be placed into Ga derived dangling bonds. These have, however, high energies causing the structure to relax from this configuration. In the relaxed structure there are developing metallic Ga-Ga bonds in favor of emptying related dangling bonds. However, our calculation gives a high surface energy. This can be understood by
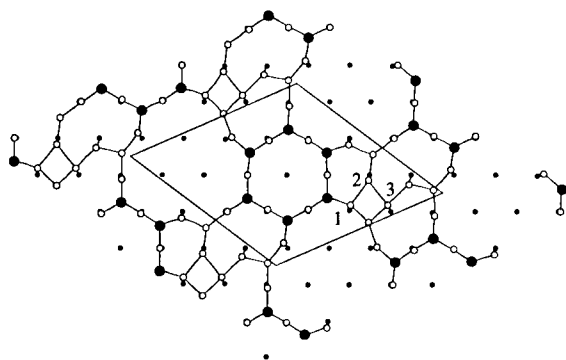
Figure 2: Top view of the Ga metallic model. Large (small) filled circles represent top (third) layer As, empty circles second layer Ga atoms.
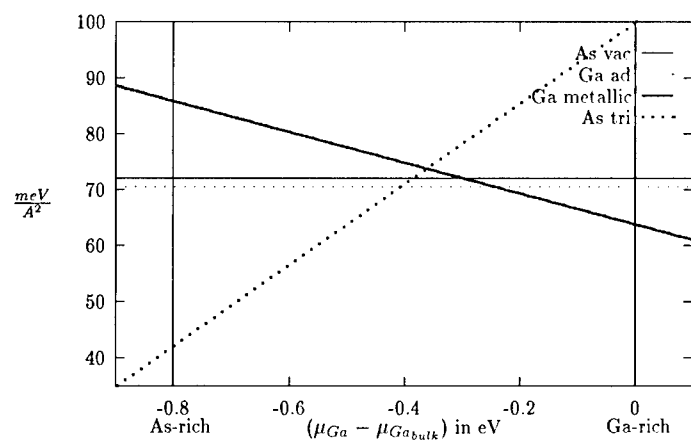
Figure 3: Surface energies of the GaAs ($\overline{1}\overline{1}\overline{1}$) structures plotted versus the Ga chemical potential. The part on the left (right) of the diagram corresponds to As (Ga) rich growth conditions.

noting that surface Ga trimers are far from being sp^2 coordinated and have thus a high surface energy.

High resolution STM images performed by Biegelsen *et al.* [47] indicate furthermore that the top layer As atoms should be arranged within a hexagonal ring. To match the STM images Biegelsen *et al.* proposed a model, where six top layer As atoms are bound to six three–fold and six two–fold coordinated Ga atoms (cf. Ref. [47]). Due to the six two–fold coordinated Ga atoms 3.75 electrons would have to occupy Ga derived dangling bonds, which strongly violates the electron counting rule. It is therefore not surprising that the atoms relax towards a configuration which fulfills the electron counting rule with only two two–fold coordinated Ga atoms and four metallic Ga–Ga bonds. The Ga–Ga bonding is achieved by forcing the Ga atoms into nearly linear Ga–Ga–As chains which are energetically very costly resulting in a high surface energy.

We then investigated the stability of a variety of different structures. The energetically most favorable configuration is shown in Fig. 2. In agreement with the previously mentioned STM pictures [47], six top layer As atoms form a hexagonal ring. Fifteen three–fold coordinated and three four–fold coordinated Ga atoms build the second layer. The three–fold coordinated Ga atoms adopt an sp^2–like hybridisation. The most important feature is the existence of the three four–fold coordinated Ga atoms, labeled 1,2,3 in Fig. 2. They exhibit weak metallic Ga–Ga bonds between them with a bond length of ≈ 2.9 Å which is comparable to the bond length in Ga bulk (≈ 2.7 Å). The structure allows an additional 2.25 electrons per $\sqrt{19} \times \sqrt{19}$ unit cell to be placed into Ga–Ga bonds before Ga derived dangling bonds would have to be occupied. All As lone pairs are filled. The electron counting rule is thus fulfilled. This and the fact that all surface Ga atoms are either sp^2 hybridised or exhibit metallic bonding and consequently do not have to adopt high energy configurations might be the reason for the low formation energy of this surface. Fig. 3 shows the energies of the most stable 2×2 reconstructions together with the proposed model for the $\sqrt{19} \times \sqrt{19}$ surface. We see that the proposed model is energetically favorable under Ga–rich growth conditions and could thus be a candidate for the observed periodicity.

Theory of threading dislocations in GaN /50/

Threading dislocations parallel to **c** and Burgers vectors **c, a** and **c+a** occur at high densities $\sim 10^9$ cm^{-2} in GaN. Recent cathodoluminescence (CL) studies of the yellow luminescence (YL) centred at ~ 2.2 eV have shown that the YL is spatially non-uniform and can be correlated with extended defects and especially low angle grain boundaries which contain dislocations [53]. On the other hand, atomic force microscopy in combination with CL has led to the conclusion that threading dislocations act as non-radiative recombination centres and degrade the luminescence efficiency in the blue light spectrum of the epilayers [54].

Adressing these experiments we use the SCC–DFTB and the *ab initio* density functional method AIMPRO [55] to investigate dislocations in GaN. Isotropic elasticity theory was used to generate the initial positions of the atoms. Relaxations were then carried out using the conjugate gradient algorithm. The dislocations are modelled in large clusters and in 576 atom ($12 \times 12 \times 1$) supercells containing a dislocation dipole in each cell in order to calculate the local line energy of the dislocations. In analogy to the formation energy for planar defects [56], we define the local line energy of extended defects as $E_{line} = \frac{1}{2L}(E - E_{bulk})$, where E is the total energy of a supercell containing two line defects, E_{bulk} is the total energy of a bulk system with the same number of atoms and $L = 5.19$ Å is the length of the supercell of the line defect in the (0001) direction. Screw dislocations in α–GaN have elementary Burgers vector $\pm$ **c** and are unusual in often being associated with nanopipes with diameters 50 - 250 Å [57]. We consider therefore a model for the open core screw dislocation ($b = \pm[000c]$). The relaxed structure (Fig. 4) preserved the hexagonal core character, demonstrating that the internal surfaces of the dislocation cores shown in (Fig. 5) are similar to (10$\bar{1}$0) facets except for the topological singularity required by a Burgers circuit. Indeed the hybridisations of the atoms situated at the wall of the open–core are very similar to the corresponding surface atoms at the (10$\bar{1}$0) surface [58]. In both cases the three fold coordinated Ga (N) atoms (no. 1 and 2 in Fig. 4,5) adopt sp^2 (p^3) like hybridisations which lower the surface energy and clean the gap [58]. However, in contrast to the (10$\bar{1}$0) surface,
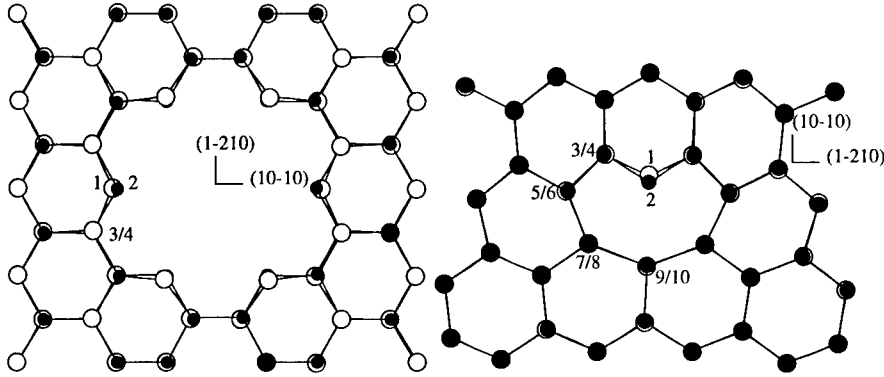
Figure 4: Top view (in [0001]) of the relaxed core of the open–core screw dislocation (b = [0001]) (*left*) and of the threading edge dislocation (b = [11$\bar{2}$0]) (*right*). The three fold coordinated atoms 1 (Ga) and 2 (N) adopt a hybridization similar to the (10$\bar{1}$0) surface atoms.

shallow gap states remain. From calculations for a full core screw dislocation and for an open core screw dislocation with a larger diameter we found an equilibrium diameter of 7.2 Å [59]. Dislocations with large open diameters that are observed might arise from kinetic factors relating to the coalescence of misaligned interfacial growth islands leading to pinholes which do not grow out [60].

Pure edge dislocations lying on {10$\bar{1}$0} planes are a dominant type of dislocation in α–GaN grown by MOCVD on (0001) sapphire, occurring at extremely high densities of $\sim 10^8 - 10^{11}$ cm^{-2}. They were modelled by relaxation of a cluster containing one dislocation with Burgers vector $a[11\bar{2}0]/3$ and a supercell with a dislocation dipole. The relaxed core is shown in Fig. 4. Again the three–fold coordinated Ga (N) atoms (no. 1 and 2 in Fig. 4) relax towards sp^2 (p^3) leading to empty Ga lone pairs pushed towards the conduction band minimum, and filled lone pairs on N atoms lying near the valence band maximum, in a manner identical to the (10$\bar{1}$0) surface. Thus threading edge dislocations are also electrically inactive.

In analogy to the open–core screw dislocations we have investigated whether the energy of the threading edge dislocation could be lowered by removing the most distorted core atoms (see Fig. 4). However, removal of either the lines of atoms 9 and 10 or the atoms 1, 2, 3, 4, 5, 6, 7, 8 and their equivalents on the right, leads to higher line energies. This implies that, in contrast with screw dislocations, the threading edge dislocations should exist with a full core. It is worth noting, however, that the core atoms 9 and 10 have very stretched bonds with bond–lengths ranging from 2.0 to 2.2 Å and thus give rise to a stress field which could act as a trap for intrinsic defects and impurities, *eg.*, the Ga vacancy O complex which could be responsible for the YL in n-type GaN [61].

VI. CONCLUSIONS

Tight-Binding as a key to complex materials properties at large scale

We have presented a straightforward extension of the standard TB theory to operation in a self-consistent charge mode based on a second order expansion of the Kohn-Sham energy functional. By this we successfully address a key problem of electronic structure theory, the development of robust, accurate, rapid, and generally transferable methods for ab-initio based simulation and characterization of large scale molecular and condensed systems. The described benchmarks
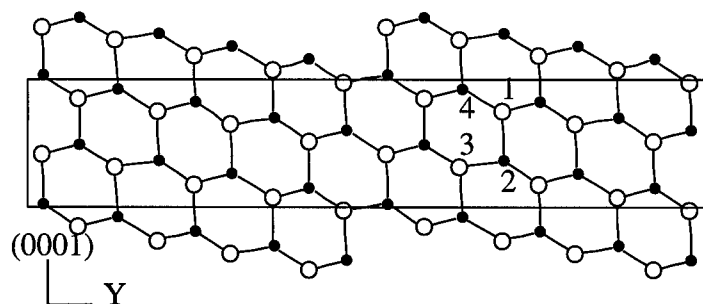
Figure 5: Projection of the wall of the open–core (d=7.2 Å) screw dislocation ($b = [0001]$). The three fold coordinated atoms 1 (Ga) and 2 (N) adopt a hybridization similar to the ($10\bar{1}0$) surface atoms.

and results clearly demonstrate the method's successful operation at sufficient accuracy on very different systems and materials, including up-to-date results for large-scale biomolecules, GaAs surface reconstructions and extended dislocations in Group-III-Nitrides. This clearly shows the usefulness of the scheme for improving various TB applications in material science.

Since further a parallelized SCC-code which extends the currently tractable system size to 1000 to 2000 atoms is already available *(M. Haugk, J. Elsner, Th. Frauenheim, to be published)* and an order-N scheme will work very soon *(M. Sternberg, G. Galli, Th. Frauenheim, to be published)* for even larger systems, the presented SCC-DFTB scheme may become one key for accurate material science simulations of ever larger systems.

REFERENCES

1. J. Harris, Phys. Rev. **31** 1770 (1985).
2. W. Foulkes, R. Haydock, Phys. Rev. B **39** 12520 (1989).
3. G. Seifert, H. Eschrig, W. Bieger, Z. Phys. Chemie *(Leipzig)* **267** 529 (1986).
4. A. P. Sutton, M. W. Finnis, D. G. Pettifor, Y. Ohata, J. Phys. C **21** 35 (1988).
5. O. F. Sankey, D. J. Niklewski, Phys. Rev. B **40** 3979 (1989).
6. D. J. Chadi, Phys. Rev. Lett. **43** 79 (1979).
7. W. A. Harrison, Phys. Rev. B **34** 2787 (1986).
8. O. K. Andersen, O. Jepsen, Phys. Rev. Letters **53** 2571 (1984).
9. E. Artacho, F. Yndurain, Phys. Rev. B **44** 6169 (1991).
10. R. E. Cohen, M. J. Mehl, D. A. Papaconstantopoulus, Phys. Rev. B **50** 14694 (1994).
11. D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert Phys. Rev. B **51** 12947 (1995).
12. F. Bechstedt, D. Reichardt, R. Enderlein, phys. stat. sol. (b) **131** 643 (1985).
13. W. A. Harrison, Phys. Rev. B **31** 2121 (1985).
14. J. A. Majewski, P. Vogl, Phys. Rev. B **35** 9666 (1987).
15. O. L. Alerhand, E. J. Mele, Phys. Rev. B **35** 5533 (1987).
16. L. Skriver, M. Rosengaard, Phys. Rev. B **43** 9538 (1991).
17. M.-H. Tsai, O. F. Sankey, J. D. Dow, Phys. Rev. B **64** 10464 (1992).
18. A. A. Demkov, J. Ortega, O. F. Sankey, M. P. Grumbach, Phys. Rev. B **52** 1618 (1995).
19. W. Kohn, L. J. Sham, Phys. Rev. **140A** 1133 (1965).

20. A. Sieck, D. Porezag, Th. Frauenheim, M.R. Pederson, K. Jackson, in print, Phys.Rev. A (Dec. 1997).

21. E.C. Honea, A. Ogura, C.A. Murray, K. Raghavachari, W.O. Sprenger, M.F. Jarrold and W.L. Brown, Nature **366**, 42 (1993).

22. R. Fournier, S.B. Sinnott and A.E. DePristo, J.Chem.Phys. **97**, 6, 4149 (1992).

23. K. Jackson, M.R. Pederson, D. Porezag, Z. Hajnal and Th. Frauenheim, Phys.Rev. B **55**, 2549 (1996).

24. M.R. Pederson und K.A. Jackson, Phys.Rev. B **41**, 7453 (1990); K.A. Jackson and M.R. Pederson, Phys.Rev. B **42**, 3276 (1990); M.R. Pederson and K.A. Jackson, Phys.Rev. B **43**, 7312 (1991).

25. P. Ordejon, D. Lebedenko and M. Menon, Phys.Rev. B **50**, 5645 (1994).

26. I. Lee, K.J. Chang and Y.H. Lee, J.Phys.: Cond.Matter **6**, 741 (1994).

27. M.V. Ramakrishna and A. Bahel, J.Chem.Phys. **104**, 24, 9833 (1996).

28. J.C. Grossman and L. Mitas, Phys.Rev.Lett. **74**, 8, 1323 (1995).

29. U. Röthlisberger, W. Andreoni and P. Giannozzi, J.Chem.Phys. **92**, 1248 (1992).

30. K.M. Ho, B.C. Pan, C.Z. Wang, J.G. Wacker, D.E. Turner and D.M. Deaven, *to be printed in* Phys.Rev.Lett. (1997).

31. F. Weich, J. Widany, Th. Frauenheim, Phys. Rev. Lett. **78** 3326 (1997).

32. A.Y. Liu & M.L. Cohen, Science **245** (1989) 841.

33. Z.J. Zhang, J. Huang, S. Fan, C.M. Lieber, Mater. Sci. and Eng. A **209** (1996) 5.

34. H. Hofsäss, C. Ronning, U. Griesmeier, Mat. Res. Soc. Symp. Proc. **354** (1995) 93.

35. Th. Frauenheim, G. Jungnickel, Th. Köhler, U. Stephan, J. Non-Cryst. Sol. **182** (1995) 186.

36. D. A. Drabold, P. A. Fedders, P. Stumm, Phys. Rev. B **49** (1994) 16415.

37. M. Elstner, D. Porezag, G. Jungnickel, Th. Frauenheim, S. Suhai, G. Seifert, submitted to PRL; and *this volume*.

38. J. Elsner, M. Haugk, D. Porezag, G. Jungnickel, Th. Frauenheim, subm. to Phys. Rev. B.

39. M. Elstner, D. Porezag, G. Jungnickel, Th. Frauenheim, S. Suhai, G. Seifert,to be published, Phys. Rev. B, and *this volume*.

40. J. Andzelm, E. Wimmer, J. Chem. Phys. **96** 1280 (1992).

41. J. S. Dewar, E. Zoebisch, E. F. Healy, J. J. P. Stewart, J. Am. Chem. Soc. **107** 3902 (1985)

42. B. Santarsiero, J. am. Chem. Soc. **112** 9416 (1990).

43. E. Tajkhorsid, M. Elstner, T. Frauenheim , S. Suhai, to be published.

44. M. Haugk, J. Elsner, Th. Frauenheim, submitted to Phys. Rev. Lett.

45. A.Y. Cho and I. Hayashi, Solid-state Electron **14**, 125 (19971).

46. R. Arthur, Surf. Sci. **43**, 449 (1974).

47. D.K. Biegelsen, R.D. Bringans, J.E. Northrup, and L.E. Schwartz, Phys. Rev. Lett. **65**, 452 (1990).

48. D. A. Woolf, D.I. Westwood and R.H. Williams, Appl. Phys. Lett. **62**, 1371 (1993).

49. E. Kaxiras, Y. Bar-Yam, J.D. Joannopoulos, and K.C. Pandey, Phys. Rev. B **35** 9636 (1987).

50. G.-X. Qian, R.M. Martin, and D. J. Chadi, Phys. Rev. B **38** 7649 (1988).

51. M.Haugk, J.Elsner and Th. Frauenheim, J. Phys. Condens. Matter **9**, 7305 (1997).

52. J. Elsner, B. Jones, P. Sitch, D. Porezag, M. Elstner, Th. Frauenheim, Phys. Rev. Lett. (27-th October 1997).

53. F. A. Ponce, D. B. Bour, W Gotz, and P. J. Wright, Appl. Phys. Lett. **68**, 57 (1996).

54. S. J. Rosner, E. C. Carr, M. J. Ludowise, G. Girlami, and H. I. Erikson, Appl. Phys. Lett. **70**, 420 (1997).

55. Jones, R.; *Phil. Trans. Roy. Soc. Lond. A* 1995, **350**, 189.

56. J. E. Northrup, J. Neugebauer, and L.T. Romano, Phys. Rev. Lett. **77**, 103 (1996).

57. W. Qian, M. Skowronski, K. Doverspike, L. B. Rowland, and D. K. Gaskill, J. Cryst. Growth, **151** 396, (1995).

58. J. E. Northrup, and J. Neugebauer, Phys. Rev. B **53**, 10477 (1996).

59. J. Elsner, R. Jones, P.K. Sitch, D. Porezag, M. Elstner, Th. Frauenheim, M.I. Heggie, S. Öberg and P.R. Briddon, Phys. Rev. Lett. **79**, 3672 (1997).

60. Z. Liliental-Weber, *et al.* to be published in the ICDS 19 proceedings (1997).

61. J. Neugebauer, and C. Van de Walle, Appl. Phys. Lett. **69**, 503 (1996).

Tight-binding calculations of electronic structure and resistivity of liquid and amorphous metals

S.K. Bose

Physics Department, Brock University, St. Catharines, Ontario L2S 3A1 CANADA

ABSTRACT

We discuss various aspects of calculating the electronic structure of liquid and amorphous metals using the recursion method and the tight-binding linear muffin-tin orbitals (TB-LMTO) basis. Resistivity calculations for such systems based on the Kubo-Greenwood formula and the TB-LMTO-recursion method are presented and compared with similar calculations based on the linear combination of atomic and atomic-like orbitals (LCAO) and the chemical pseudopotential approach. Results for amorphous Fe and Co and liquid Hg, Pd, and some 3d transition metals are presented. Sources of error in the calculation and ways to improve upon the present calculations are discussed.

INTRODUCTION

In this work we will discuss the calculation of electronic structure and the electrical conductivities of liquid and amorphous metals using the tight-binding linear muffin-tin orbitals[1,2] basis. There are obvious motivations for such calculations. Numerical computation of the electronic structure of large disordered models, based on real space methods[3], requires short ranged basis functions. At present the tight-binding linear muffin-tin orbitals (TB-LMTO) basis is the only such minimal basis available, whose properties are determined by the density-functional theory and which thus provides the way for a self-consistent first-principles calculation. A TB-LMTO basis orbital, centered on one atom, depends sensitively via the potential parameters and the TB structure constants[1,2] on the chemical nature and spatial distribution, respectively, of its neighbors. This property is in marked contrast with the conventional TB bases, whose properties are independent of the local environment. A TB-LMTO basis is thus a natural choice for calculations involving amorphous systems. The TB-LMTO-recursion method has been used with considerable success to study the electronic structure and related properties of metallic glasses[4-7], liquid metals[8,9], quasicrystals[10,11] and amorphous semiconductors[12]. Often the TB-LMTO-recursion method provides a level of accuracy that is comparable to that of the LMTO-supercell method[4].

LMTO and TB-LMTO

For details of the LMTO and the TB-LMTO method we direct the readers to Refs.[1,2]. In a nearly orthogonal LMTO representation, the Hamiltonian matrix elements, under the atomic sphere approximation, can be written as

$$H_{RL,R'L'} = E_{\nu Rl}\delta_{RR'}\delta_{LL'} + h^0_{RL,R'L'} = C_{RL}\delta_{RR'}\delta_{LL'} + \Delta^{1/2}_{RL}\left[S^0(1-\gamma S^0)^{-1}\right]_{RL,R'L'}\Delta^{1/2}_{R'L'} .$$

$$(1)$$

Here R denotes the centers of the muffin-tin spheres into which the entire space in the solid or the liquid is divided, and $L(\equiv l, m)$ is the collective angular momentum index. $E_{\nu Rl}$ denote the reference energies, usually chosen as the centers of the occupied parts of the respective l-bands. The properties of the atoms enter the Hamiltonian matrix H via the matrices $X(X = C, \Delta, \gamma)$, which are diagonal in the indices $\vec{R}$ and L and are independent of m, i.e.,

$$(X)_{RL,R'L'} = X_{RL}\delta_{RR'}\delta_{LL'} = X_{Rl}\delta_{RR'}\delta_{LL'} .$$

Also $X_{Rl} = X_l^Q$, where Q is the type of the atom occupying the site $\vec{R}$. The matrix elements X_{RL}, called the potential parameters, can be derived from the knowledge of the functions $\Phi_{\vec{R}l}$ and their energy derivative at the reference energy $E_{\nu Rl}$, $\Phi_{\vec{R}l}$ being the normalized solution of the radial part of the wave equation for orbital angular momentum l inside the sphere at $\vec{R}$ for the reference energy $E_{\nu Rl}$. S^0 is the canonical structure constant matrix, which depends on the relative positions of the sites $\vec{R}$, but is independent of the type of atoms that occupy the sites.

The envelope functions of the LMTOs giving rise to the Hamiltonian of Eq.(1) are long-ranged. In the TB-LMTO method one uses short-ranged envelope functions. A diagonal "screening matrix", $\alpha_{RL,R'L'} = \alpha_{RL}\delta_{RR'}\delta_{LL'}$, is introduced in the theory. The resulting screened or short-ranged envelope functions give rise to short-ranged basis functions and short-ranged Hamiltonian matrix elements with the same form as in Eq.(1):

$$H_{RL,R'L'}^\alpha = C_{RL}^\alpha \delta_{RR'}\delta_{LL'} + (\Delta_{RL}^\alpha)^{1/2}\left[S^0(1 - \alpha S^0)^{-1}\right]_{RL,R'L'}(\Delta_{R'L'}^\alpha)^{1/2} . \tag{2}$$

For disordered structures matrix elements of the screened structure constants $S^\alpha = S^0(1 - \alpha S^0)^{-1}$ connecting atoms further apart than the distance corresponding to the first minimum in the pair-distribution function can be neglected without serious error. The relation between the LMTO Hamiltonian H given by Eq.(1) and the TB-LMTO Hamiltonian, $H^\alpha = E_\nu + h^\alpha$, is given by

$$H = E_\nu + h^\alpha(1 + o^\alpha h^\alpha)^{-1} = H^\alpha - h^\alpha o^\alpha h^\alpha + h^\alpha o^\alpha h^\alpha o^\alpha h^\alpha - ...... , \tag{3}$$

where the parameter o^α determines the nonorthogonality of the TB-LMTO basis with the corresponding overlap matrix being given by $O^\alpha = (1 + h^\alpha o^\alpha)(1 + o^\alpha h^\alpha)$. The relationship between the TB-LMTO potential parameters, with superscript α, appearing in Eqs.(2) and (3) and the standard LMTO potential parameters appearing in Eq.(1) can be found in Ref.[2]. It is clear from Eq.(3) that the accuracy of the TB-LMTO Hamiltonian is crucially dependent on the parameter o^α which determines the degree of nonorthogonality of the TB-LMTO basis. The parameters $(o^\alpha)^{-1}$ have the dimensions of energy and provide a measure of the energy window about the reference energies E_ν for which the DOS results obtained from H^α are reliable. Solving the eigenvalue problem for H^α only is equivalent to neglecting the nonorthogonality of the TB-LMTO basis. This neglect is justifiable if all the o^α parameters are small. Since relevant bandwidths are of the order of 1 Ry, o^α should be less than 1 Ry^{-1}, or $(o^\alpha)^{-1}$ should be at least 1 Ry, preferably higher. In section 3 we will illustrate the effect of neglecting the nonorthogonality of the TB-LMTO basis by comparing the densities of states (DOS) calculated by using the Hamiltonians $H_1 = H^\alpha$ and $H_2 = H^\alpha - h^\alpha o^\alpha h^\alpha$ for liquid Hg and Pd. Some examples are also discussed in Refs.[4,13].

Pair distribution function in liquid Pd at 1580°C

Fig.1. Pair distribution function of a simulated 1372 atom cluster compared with the experimental pair distribution function in liquid Pd at 1580°C.

SIMULATION OF LIQUID AND AMORPHOUS CLUSTERS

Clusters representing the liquid and amorphous metals to be discussed in this work were all generated via Monte Carlo(MC) simulation using the Metropolis Scheme[14]. Results for liquid Pd and Hg are reported here for the first time. Therefore we provide some details regarding the simulation of these two systems. For simulation details involving other liquid and amorphous metals, whose results are occasionally mentioned for the sake of comparison and the completeness of the discussions, we refer the readers to previous work (liquid 3-d transition metals[8], liquid La[9], amorphous Fe and Co[15]).

In order to generate liquid Pd clusters we employed the two body potential used by Gillan[16] in connection with the study of hydrogen diffusion in Pd. This potential is based on a nearest neighbor model developed by van Heugten[17] and co-workers by fitting to the phonon spectra. It is in the form of a sixth-degree polynomial and has a cutoff at 3.305Å, which lies between the first and second neighbors in fcc Pd. The minimum in the potential is at 2.884Å. The depth is 0.1607 eV or 2138 K. Several(10-20) 1372-atom clusters were generated using this two-body potential. The number densities of the clusters were chosen to be 0.0594Å^{-3}, the density of liquid Pd at 1580°C[18]. The pair distribution functions of all such clusters, in equilibrium, compared well with the experimentally determined pair distribution function[18]. The pair distribution function of one such simulated 1372-atom cluster is shown in Fig.1. Except for a small difference in the location of the peak, the agreement is excellent.

Since there is no suitable pair potential available in the literature for solid or liquid Hg, we attempted to construct one based on a method proposed by Carlsson et al.[19]. These authors derived a procedure for inverting the cohesive energy of an isostructural one-component system as a function of volume in order to obtain a pair potential. Essentially, the cohesive energy is written as a lattice sum of some effective two body potential. The series for the cohesive energy is inverted to obtain the pair potential in the form of a series involving the cohesive energies for various lattice parameters. We used the cohesive energy obtained by standard LMTO-ASA method for Hg in the fcc structure. By writing the cohesive energy as a sum of two body potentials involving more and more distant neighbor shells one can hope to converge to a potential. We found that for practical

purposes of generating a liquid cluster there is no obvious advantage in trying to achieve such a convergence. For example, the potential obtained with the von Barth-Hedin[20] exchange-correlation potential and with eight neighbor shells showed little difference from that obtained with only two neighbor shells. However, in the MC simulation, for reasons of computer time involved, the potential is cut off beyond distances larger than the second peak in the pair distribution function. The main features of the potential, such as the depth and the location of the minimum, are converged within 5% with only just two shells of neighbor. Thus we used a two body potential obtained by writing the cohesive energy as a sum of interactions involving only two shells of neighbors in the fcc structure. For use in the MC simulation we fitted it to the Morse function form:

$$\phi(r) = C(e^{-2\alpha(r-r_0)} - De^{-\alpha(r-r_0)}), \tag{4}$$

with $C = 0.18343$ eV, $\alpha = 1.8537 \text{Å}^{-1}$, $r_0 = 3.1366 \text{Å}$, $D = 1.6877$. The potential was cut off beyond 4.5Å. As discussed in detail by Carlsson et $al.$[19], the depth of the potential thus obtained is too large compared with the potentials obtained by the pseudopotential approach (for sp-bonded metals) or by fitting to experimental cohesive energy and other elastic properties. For example, the potential given by Eq. (4) has a depth of 0.128 eV, whereas the melting point of Hg is 233 K or 0.02 eV. The simulation using the above potential had to be carried out at a very high temperature (above 3000 K) to obtain pair distribution functions that match the experimental curves in the temperature range [150°C - 250°C][18]. In the context of the present work the simulation temperature has no physical significance beyond simply a means of generating liquid clusters via a canonical MC simulation. In Fig. 2 we compare the pair distribution functions of 1372 atom clusters obtained via simulations with the experimental distributions at 150°C and 250°C. The number densities of these clusters are chosen to be the appropriate values for the corresponding temperatures (0.0386Å^{-3} at 250°C, and 0.0397Å^{-3} at 150°C[18]). The agreement between the experimental and the simulated pair distribution functions is excellent at 250°C and becomes poorer for lower temperatures. However, for 150°C the agreement is still good enough to permit the use of the simulated clusters as appropriate models for liquid Hg at 150°C. Potentials constructed via nonlocal exchange- correlation functional (e.g Perdew-Wang[21]) differ somewhat from those obtained with local functionals in terms of the depth and the location of the minimum. However, the clusters obtained via these potentials show similar pair distribution functions. Experimental pair distribution functions of liquid Hg for temperatures higher than 250°C are not available. We were thus unable to generate models appropriate for higher temperatures. The details of derivation of the potentials and the simulations will be published elsewhere.

CALCULATED ELECTRONIC STRUCTURE OF LIQUID AND AMORPHOUS METALS

TB-LMTO studies of the electronic structure of liquid 3d transition metals[8], amorphous Fe and Co[15], and liquid La[9] have been reported previously. The double peaked structure of the DOS in the crystalline phases of the 3d transition metals is found to survive in both the amorphous and the liquid phases, with small (less than 5%) band narrowing occurring between the crystalline, amorphous and the liquid phases as a result of decreasing density. The agreement with experimental results (such as photoemission, magnetic moment for the amorphous phase), whenever available, is good. In the following we present the electronic
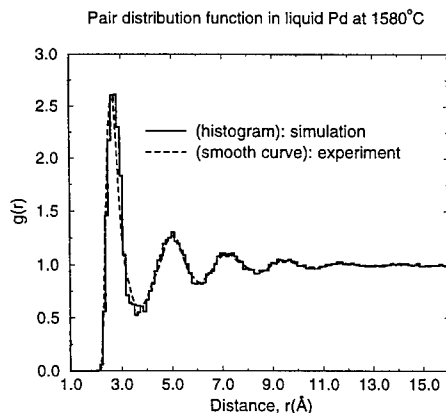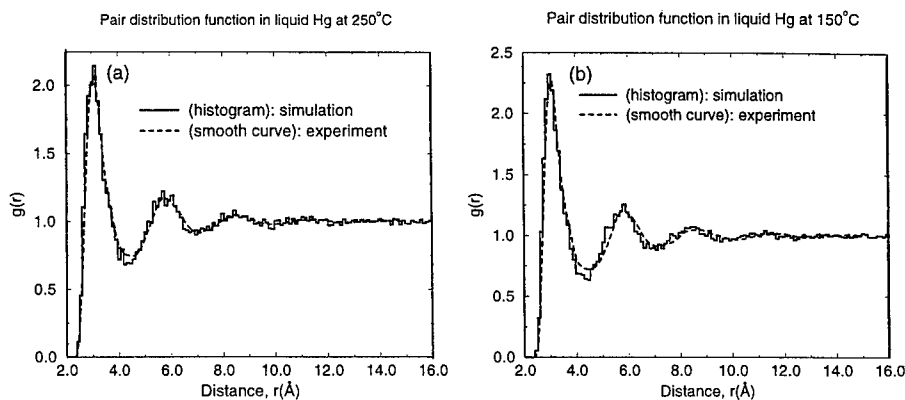
Fig.2. Pair distribution function of simulated 1372 atom Hg clusters compared with the experimental pair distribution function in liquid Hg at (a) 250°C and (b) 150°C.

structures of liquid Pd and Hg, which have not been studied so far using the tight-binding method.

Electronic properties of liquid mercury seem unusual compared with most other liquid metals. It has a rather high dc resistivity, which decreases on addition of most metallic impurities. In order to explain this and several other empirical facts Mott[22] proposed the pseudogap model, suggesting a substantially reduced density of states at the Fermi energy. Since then several authors have tried to compute the electronic density of states in liquid Hg using various methods. Some of these have been reviewed by Ballentine[23]. Mattheiss and Warren, Jr.[24] presented a band model calculation for the electronic structure of expanded liquid mercury. Yonezawa *et al.*[25] employed an effective medium theory and the single-site approximation to discuss the electronic structure of some liquid metals including liquid Hg. The DOS calculations of Ballentine and co-workers (see ref. 23) are based on the Green's function approach, with the self-energy calculated by using a nonlocal pseudopotential. Most other calculations for liquid Hg are based on various other model pseudopotentials, with their suitability being debated among the authors. As pointed out by Ballentine[23], in spite of all these calculations and the available experimental results, there remain open questions about the electronic structure of liquid mercury. In this paper the first density functional(TB-LMTO) calculation for liquid mercury based on a realistic structural model is presented. In Fig.3(a) we show the DOS's calculated for fcc Hg at a density of 0.0386 atoms/$Å^3$, the density of liquid Hg at 250°C, using the standard LMTO method and the recursion method[3] applied to a 1372 atom fcc cluster with periodic boundary conditions. The Hamiltonian H_2, as defined at the end of section II, was used for the recursion method calculation and the continued fraction was terminated using the linear predictor method of Allan[26]. The differences in the band widths in the two calculations is due to the differences in the Hamiltonians used, H_2 being only the first two terms in the infinite series on the r.h.s. of Eq.(3) representing the full LMTO-ASA Hamiltonian.

Both DOS's show a pseudogap with the Fermi energy (E_F) lying slightly above the minimum of the pseudogap. The similarity of the two curves shows the reliability of the real space method we are going to use for calculating the liquid state electronic structure and

109
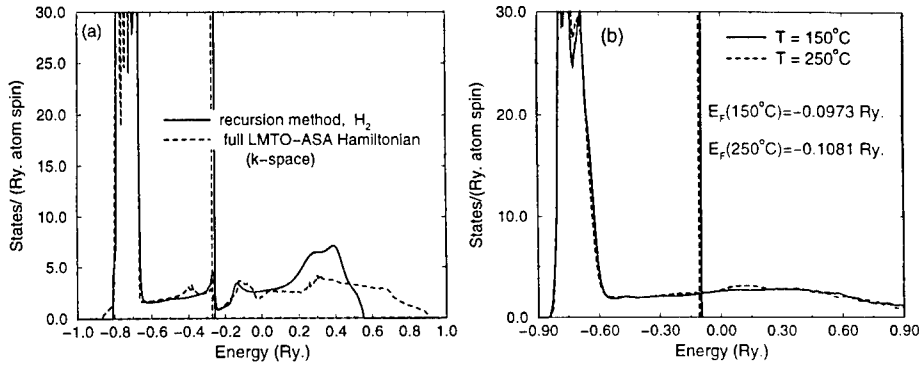
Fig.3. (a) The DOS of fcc Hg at the density of liquid Hg at 250°C, calculated via standard LMTO-ASA (k-space) method and the rescursion method using Hamiltonian H_2, (b) the DOS in liquid Hg at 250°C and 150°C calculated via recursion method using Hamiltonian H_2. Results show 100 atom averages.

also that features such as a deep minimum are reproducible via the recursion method[3]. The DOS in the ground state of Hg (rhombohedral) does not show such a pseudogap. However, the DOS in the fcc structure at the density of solid Hg (which is only marginally higher in energy with respect to the rhombohedral structure) does have a minimum close to E_F. This pseudogap, related to the separation of the centers of the $'6s'$ and $'6p'$ bands, is expected to increase with increasing temperature or decreasing density, ultimately producing a metal-semiconductor transition. However, as shown in Fig.3(b), the pseudogap does not exist in the liquid state around the temperatures 150-250°C. In Fig.3(b) we show the DOSs for 1372 atom clusters representing liquid Hg at 250 and 150°C, obtained using the Hamiltonian H_2. Each DOS shown is an average over the local DOSs for 100 atoms in the cluster. The pseudogap observed for the crystalline phase disappears due to fluctuations in local environments causing local $'6s'$ and $'6p'$ DOSs to vary considerably in their peaks heights, centers and widths. These variations are sufficient to wipe out any pseudogap reminiscent of the crystalline phase. Thus although it is expected that at higher temperatures this pseudogap should reappear and become deeper, eventually giving rise to the semiconducting state, it is nonexistent at low temperature (below $\sim$ 300°C) liquid mercury and cannot be responsible for the observed decrease in its resistivity on alloying with other metals. In Fig.4(a) we show the s- and p-orbital projected DOSs in liquid Hg at 250°C. In Fig.4(b) we compare the s- and p-orbital projected DOSs in liquid Hg clusters at 250 and 150°C. It is apparent that there is no pseudogap at the Fermi energy. The deep minimum occurring in the s,p- DOS at energies below E_F is not due to increased separation between the centers of the 6s and 6p bands, but due to the hybridization of the d-band with the sp-band, sometimes referred to as a 'Fano effect'. As shown in Fig.3(b), apart from an almost imperceptible decrease in the d-bandwidth there is virtually no change in the DOS as the temperature increases from 150°C to 250°C, the DOS at E_F remaining around [2-2.5] states/(Ry. atom), the same as in the rhombohedral ground state of Hg. Both s- and p-DOSs are found to remain unchanged in our calculation as the temperature increases from 150°C to 250°C.
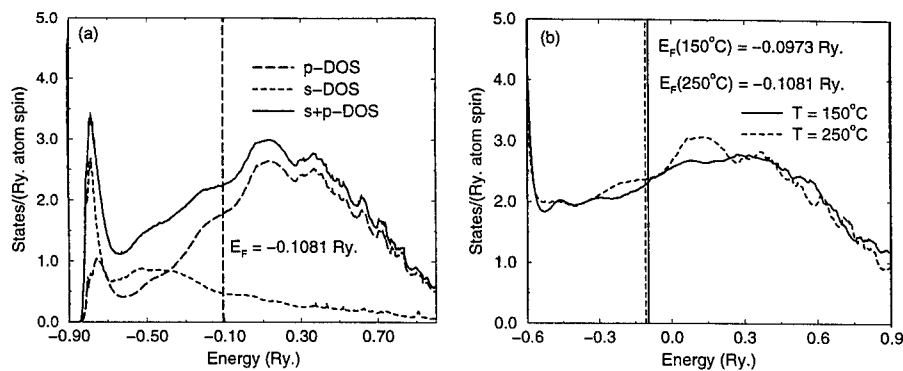
Fig.4. (a) s- and p-orbital projected DOSs in liquid Hg at 250°C, (b) total s+p DOSs in liquid Hg at 250°C and 150°C. Vertical lines show the locations of the Fermi energy.
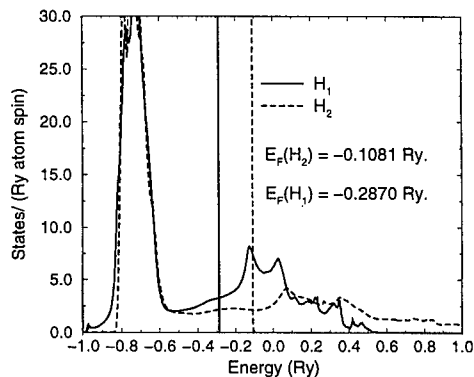
Fig.5. A comparison of the DOSs in liquid Hg at 250°C, calculated by using the Hamiltonians H_2 and H_1. Each DOS is a 10 atom average in a 1372 atom cluster. Vertical lines show the locations of the Fermi energy.

The lack of variation in the s-DOS is in agreeemnt with the measurement of the Knight shift in liquid Hg, which is found to be more or less independent of the density around these values[24]. The effect of the nonorthogonality of the TB-LMTO basis on the DOS is shown in Fig.5 by comparing the DOSs calculated with the Hamiltonians H_2 and H_1 for a 1372 atom cluster representing liquid Hg at 250°C. The effect of the second term on the r.h.s of Eq.(3) is to shift most energy eigenvalues to higher energy, increasing the bandwidth as well as pushing the bottom of the band to higher energy compared with the TB-LMTO Hamiltonian H_1.

We have computed the electronic density of states in liquid Pd using the TB-LMTO method and the structural model presented in Fig.1. Several 1372- atom clusters representing liquid Pd at 1580°C were used to compute an average DOS, and in each case an average over the local DOSs corresponding to 100 atoms were considered. In Fig.6 we show the DOSs for liquid Pd at 1580°C calculated by using Hamiltonians H_1 and H_2. For the sake of comparison we also show the corresponding DOSs calculated for 1372 atom fcc Pd

DOS in liquid Pd at 1580°C

(a)

States/(Ry. atom spin)

16.0
12.0
8.0
4.0
0.0

$E_F(H_1)$ = -0.1905 Ry.
$E_F(H_2)$ = -0.1894 Ry.

—— H_1-DOS
----- H_2-DOS

-0.80   -0.40   0.00   0.40   0.80

Energy (Ry.)

DOS in fcc Pd at the density of liquid Pd at 1580 °C
recursion method

(b)

States/(Ry. atom spin)

25.0
20.0
15.0
10.0
5.0
0.0

—— H_2
----- H_1

$E_F(H_1)$ = -0.2205 Ry.
$E_F(H_2)$ = -0.2239 Ry.

-0.8 -0.6 -0.4 -0.2 0.0 0.2 0.4 0.6 0.8 1.0
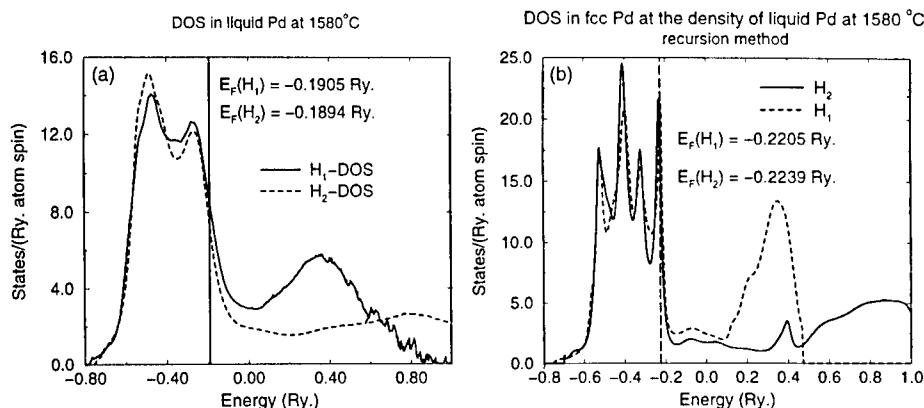
Energy (Ry.)

Fig.6. (a) DOSs in liquid Pd at 1580°C calculated by using the Hamiltonians H_1 and H_2 and (b) the same for fcc Pd at the same density as the liquid cluster. Vertical lines show the locations of the Fermi energy.

clusters at the same number density(0.0594 atoms $\mathring{A}^{-3}$) as the liquid cluster. Here the effect of the nonorthogonality of the TB-LMTO basis seems negligible for the occupied part of the DOS. The total bandwidth of the Hamiltonian H_1 is smaller. The correction term hoh in H_2 pushes the energy eigenvalues on the high energy side to higher values, while the DOS for lower energies remains nearly unaffected. The number density of crystalline fcc Pd is 0.068 atoms $\mathring{A}^{-3}$, about 13% higher than the liquid at 1580°C. For the fcc structure the total d-bandwidth decreases by about 20% (from 0.5 Ry. to 0.4 Ry.) as the density decreases from 0.068 atoms $\mathring{A}^{-3}$ to 0.0594 atoms $\mathring{A}^{-3}$, the density of liquid Pd. However, in the liquid state, because of the fluctuations in local environment, the total d-bandwidth increases to 0.55 Ry., 10% higher than the ground state fcc value (although the root mean square d-bandwidth remains slightly smaller). The DOS at E_F is about 17 states/(Ry. atom) in the liquid state, compared with about 27 states/(Ry. atom) in the fcc ground state. The shape of the DOS is similar to that of 3d liquid transition metals. The computed DOS compares well with the LMTO supercell calculation of Jank et al[27]. However, there are differences in the structural models used. The agreement with the experimental pair distribution function is slightly better for our model around the first peak of the pair distribution function, and considerably more so around the second and third peaks and beyond.

ELECTRICAL CONDUCTIVITY

In this section we discuss how the TB-LMTO method can be used to carry out a real space calculation of the electrical conductivity of disordered metals and alloys. Interested readers are also urged to consult the work by D. Mayou[28] and D. Mayou and Khanna[29], where alternative approaches to this problem, based on the properties of orthogonal polynomials, are presented.

According to the Kubo-Greenwood[30,31] formula the diagonal elements of the zero-temperature dc conductivity tensor in the eigenfunction representation are given by

$$\sigma_{jj} = \frac{e^2 h}{\Omega} \sum_{m,n} | < E_m|v_j|E_n > |^2 \delta(E_m - E_F)\delta((E_n - E_F)) , \tag{5}$$

where Ω is the sample volume, h is the Planck's constant, E_F is the Fermi energy and v_j is the j component of the velocity operator. Relating the Dirac-δ function to the imaginary part of the Green's function and using the expression

$$\sum_m \delta(E - E_m) < E_m|f(E)|E_m >= g(E)\overline{f(E)}_{E_m=E} , \tag{6}$$

where the bar implies an average over the eigenfunctions with energy E, and $g(E)$ is the sample DOS at energy E, Eq.(13) can be given a physically transparent form:

$$\sigma_{jj} = \frac{e^2}{\Omega_a} n(E_F)D(E_F) . \tag{7}$$

Here Ω_a is the volume per atom, $n(E_F)$ is the DOS per atom at the Fermi energy and $D(E_F)$ is the diffusivity given by

$$D(E_f) = -\hbar \lim_{\epsilon \to 0^+} Im \overline{[< E_m|v_j G(E_f + i\epsilon)v_j|E_m >]}_{E_m=E_F} . \tag{8}$$

Apart from a numerical factor, $D(E_F)$ can be considered as the average local DOS projected onto the states $v_j|E_m >$, and can be calculated via the recursion method[3]. The eigenvectors $|E_m >$ can be calculated by a filtering technique used originally by Kramer and Weaire[32]. The matrix elements of the velocity operator in the TB-LMTO basis are

$$(v_j)_{\beta\gamma} = (\frac{i}{\hbar}) \sum_\delta (H_{\beta\delta}x^j_{\delta\gamma} - x^j_{\beta\delta}H_{\delta\gamma}) , \tag{9}$$

where x and H denote the position and the Hamiltonian operators and the subscripts denote the combined angular momentum and the site indices. The matrix elements of the position operator, $x^j_{\gamma\beta}$, can be written as

$$x^j_{\gamma\beta} =< \chi^\alpha_\gamma|x^j|\chi^\alpha_\beta >= \frac{1}{2}(x^j_\gamma + x^j_\beta)O^\alpha_{\gamma\beta}+ < \chi^\alpha_\gamma| \left[x^j - \frac{1}{2}(x^j_\gamma + x^j_\beta)\right] |\chi^\alpha_\beta > , \tag{10}$$

where x^j_β is the x^j coordinate of the atomic nucleus on which the orbital β is centered. $|\chi^\alpha_\beta >$ represents a TB-LMTO orbital with the subscript denoting jointly a site and the collective angular momentum index L. The second term on the right hand side of the above equation, the so-called dipole term, is neglected. If the diffusivity calculation is carried out by using the Hamiltonian H_2, then the nonorthogonality of the basis is small and can be neglected. For calculation involving H_1 the overlap matrix $O^\alpha_{\gamma\beta}$ must be included in the computation of the matrix elements of the position operator. Further details of the calculations of these matrix elements appear in Ref.[8].

The resistivity of amorphous iron, with the diffusivity calculated via the first order Hamiltonian H_1, comes out to be 125 $\mu\Omega$cm. This value lies intermediately between the reported measurements: $100\pm20\mu\Omega$cm[33] and $150\mu\Omega$cm[34]. The difference between the two measured resistivities of amorphous Fe films is not necessarily to be ascribed to experimental error, since there is no evidence that the structures of the amorphous films were

the same in the two experiments. For α-Co, a similar calculation based on the most tight-binding Hamiltonian H_1 yields a resistivity value 110 $\mu\Omega$cm, a bit lower than the measured resistivity 115 $\mu\Omega$cm of liquid Co[35,36].

In ref.[8] we have shown that the above TB-LMTO-recursion method, based on the first order Hamiltonian H_1, yields resistivity values in good agreement with experiments for liquid $3d$ transition metals Cr, Mn, Fe and Co. For liquid Ni the calculated value was twice as large as the measured value. The reasons behind large errors for liquid Ni, practical limitations of the method and the ways of improvement have been suggested in ref.[8]. The accuracy obtained via H_1 is sufficient for liquid Cr, Mn, Fe and Co, but not for liquid Ni. The parameter o^α for the d-states steadily increases across the transition metals series, from 0.3 Ry^{-1} for liquid Cr to 1.35 Ry^{-1} for liquid Ni. Thus for liquid Ni the correction due to the term $h^\alpha o^\alpha h^\alpha$ is significant around the Fermi energy. One way to correct the situation is to perform the DOS calculation with the reference energies E_ν put at the centers of the occupied parts of the partial l-bands, but the diffusivity calculation(via H_1) with all the E_ν's set to E_F. The choice of the E_ν equal to the centers of the occupied parts of the partial l-bands ensures that the DOS at E_F and E_F itself are calculated properly. The choice of E_ν being equal to E_F should then make the diffusivity calculation most accurate, since the latter involves only the states near E_F. This prescription for calculating the diffusivity with H_1 yields resistivity value for liquid Ni in agreement with the measured value[36].

The resistivity calculation for the clusters representing liquid Pd and liquid Hg, discussed in the earlier sections, was carried out using the Hamiltonian H_2. The calculated value for liquid Hg at 250°C was approximately 200 $\mu\Omega$cm. The measured resistivity values at 100°C and 500°C are 103.3 $\mu\Omega$cm and 160 $\mu\Omega$cm, respectively. A similar diffusivity calculation using H_2 yields a resitivity of 120 $\mu\Omega$cm for liquid Pd at 1580°C, while the experimental value is 85 $\mu\Omega$cm. The agreement with the measured value can perhaps be improved using the Hamiltonian H_1, with the reference energies E_ν being replaced by E_F for the diffusivity calculation. This calculation has not yet been performed.

By resolving the eigenvectors at E_F into their orbital components we can estimate the contribution to the resistivity purely from the diagonal (s, p. and d) channels as well as the contributions from the mixed (off-diagonal) channels. We denote by D_s, D_p, and D_d the contributions of s, p, and d orbitals. respectively, to the full diffusivity, divided by the fractional s, p, and d components of the total DOS. For liquid and amorphous 3d metals both D_s and D_p are found to be 7 to 12 times larger than D_d (for calculations involving H_1). However, conductivity is still dominated by the d states in the sense that σ_{dd} is approximately five to six times larger than σ_{ss} and σ_{pp}, because of substantially large weight of the d states at E_F. The off-diagonal channels contribute significantly to the conductivity. The off-diagonal contribution is usually negative, and is anywhere between 50% and 100% of the diagonal contribution. The hybridization between the s, p, and d bands thus impedes the process of conduction. This result is in marked variation with that of similar calculations[37] employing LCAO (linear combination of atomic-like orbitals) basis and the chemical pseudopotential scheme (see article by Bullett in Ref. [3]), where the off-diagonal channels were found to have negligible contribution to the process of conduction, inspite of substantial s-d hybridization in the Hamiltonian. These LCAO-based calculations[37(a)] had produced resistivity results in reasonable agreement with measured values for liquid V, Cr, Mn and Fe. But the results for liquid Co and Ni were poor, with liquid Ni being the most troublesome case[37(b)].

The most noteworthy source of error in the above calculations is approximate treatment

of the velocity matrix elements. The neglect of the dipole trem in Eq.(10) can be justified if the basis consists of only s- and d-orbitals, with pure s- and d-symmetry. The inclusion of the p-orbitals in the basis, as well as the fact that the TB-LMTO orbitals lack pure angular momentum (L) symmetry[1,2], makes the validity of this approximation questionable. Apart from this, there are possible errors associated with the recursion method, and inadequacies in structural models. Inspite of these drawbacks the method remains promising for high resistivity liquid and amorphous metals, being one of the very few methods available at present for a first-principles study of the transport properties of such systems.

COMMENTS AND CONCLUSIONS

In the preceding sections we have discussed how the LMTO method can be used to study electronic structure and transport properties of liquid and amorphous metals. The method can be used at various levels of sophistication, ranging from the level of semi-empirical tight-binding schemes to the level of the most elaborate charge self-consistent first principles calculation. The errors involved at various levels of approximation used to implement the method are well-understood and can be removed with increased computational efforts. Some recent developments[38] within the LMTO scheme provide an improved basis set for describing the electronic structure of inhomogeneous matter. These involve screened spherical waves with position dependent screening parameters. The errors in the eigenvalues obtained via this method are proportional to fourth order in deviations from the reference energies E_ν. This screened spherical wave basis has not yet been used for liquids and amorphous systems , but certainly holds much promise.

ACKNOWLEDGEMENT

REFERENCES

1. O.K. Andersen and O. Jepsen, Phys. Rev. Lett.**53**, 2571 (1984).
2. (a) O.K. Andersen, O. Jepsen, M. Sob in *Electronic structure and its applications*, edited by M.Yossouff, Lecture Notes in Physics, vol.283, Springer-Verlag, Berlin, 1987, pp.1-57; (b) O.K. Andersen, O. Jepsen and D. Glötzel in *Highlights of Condensed Matter Theory*, edited by F. Bassani, F. Fumi, and M.P. Tosi, North-Holland, Amsterdam, 1985, pp.59-176.
3. R. Haydock, Solid State Physics: Advances in Research and Applications, vol.35, edited by F. Seitz and D. Turnbull, Academic Press, New York, 1980, pp.215-94; see also articles by V. Heine and M.J. Kelly in the same volume.
4. S.K. Bose, S.S. Jaswal, O.K. Andersen and J. Hafner, Phys. Rev. B**37**, 9955 (1988).
5. H.J. Nowak, O.K. Andersen, T. Fujiwara, O. Jepsen, and P. Vargas, Phys. Rev. B **44**, 3577 (1991).
6. I. Turek, J. Phys.: Condens. Matter**2**, 10559 (1990).
7. I. Turek, J. Mag. & Mag. Materials **98**, 119 (1991).
8. S.K. Bose, O. Jepsen and O.K. Andersen: Phys. Rev. B **48**, 4265 (1993).

9. S.K. Bose, O. Jepsen and O.K. Andersen, J. Phys.: Condens. Matter **6**, 2145 (1994).

10. J. Hafner and M. Krajči, Phys. Rev. Lett. **68**, 2321 (1992).

11. R.F. Sabiryanov, S.K. Bose and S.E. Burkov, J. Phys.: Condens. Matter **7**, 5437 (1995).

12. S.K. Bose, K. Weiner, and O.K. Andersen, Phys. Rev. B **37**, 6262 (1988).

13. S.K. Bose, J. Kudronovský, I.I. Mazin and O.K. Andersen, Phys. Rev. B**41**, 7988 (1990).

14. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J.Chem.Phys **21**, 1087 (1953).

15. R.F. Sabiryanov, S.K. Bose and O.N. Mryasov, Phys. Rev. B**51**,8958(1995).

16. M.J. Gillan, J. Phys. C: Solid State Phys.**19**, 6169 (1986).

17. W.F.W.M van Heugten, Phys. Stat. Sol. B**86**, 277(1978); W.F.W.M van Heugten and J.Th.M de Hosson, Phys. Stat. Sol. B**90**, 225(1978)

18. Y. Waseda,*The Structure of Non-Crystalline Materials*, Appendix 8 (McGraw-Hill, New York, 1980).

19. A.E. Carlsson, C.D. Gelatt, Jr., H. Ehrenreich, Phil. Mag. A**41**, 241 (1980).

20. U. von Barth and L. Hedin, J. Phys. C: Solid State Physics**5**, 1629(1972).

21. J. P. Perdew and Y. Wang, Phys. Rev. B**33**, 8800 (1986).

22. N.F. Mott, Phil. Mag.,**13**, 989 (1966).

23. L.E. Ballentine, Adv. Chem. Phys., vol. 31, 263 (1975).

24. L.F. Mattheiss, and W.W. Warren,Jr., Phys. Rev. B**16**, 624 (1977).

25. F. Yonezawa and F. Martino, Sol. State Commun.**18**, 1471 (1976); F. Yonezawa, F. Martino and S. Asano, in *Liquid Metals 1976*, edited by R. Evans and D.A. Greenwood, (Institute of Physics, London, 1977).

26. G. Allan, J. Phys. C: Solid State Phys.**17**, 3945 (1984); and in *The Recursion Method and its Applications*, edited by D.G. Pettifor and D.L. Weaire, Springer Series in Solid State Sciences Vol. 58 (Springer-Verlag, Berlin, 1985),p.61.

27. W. Jank, Ch. Hausleitner and J. Hafner, J. Phys.: Condens. Matter **3**, 4477 (1991).

28. D. Mayou, Europhys. Lett. **6**, 549 (1988).

29. D. Mayou and S.N. Khanna, J. Phys. I France **5**, 1199 (1995).

30. R. Kubo, Can. J. Phys. **34**, 1274 (1956).

31. D. Greenwood, Proc. Phys. Soc. London **71**, 585 (1958).

32. B. Kramer and D. Weaire, J. Phys. C: Solid State Phys. **11**, L5 (1978).

33. R.V. Aldridge and S. J. Raeburn, Phys. Lett.**56A**, 211 (1976).

34. J.G. Wright, *Liquid Metals* (Institute of Physics Conference Series, vol.30), 251 (1977).

35. Y. Kita and Z. Morita, J. Non-Cryst. Solids **61**& **62**, 1079 (1984).

36. Metals Reference Book, 5th ed., edited by Colin J. Smithells (Butterworths, London, 1976), p. 947.

37. (a) L.E. Ballentine, S.K. Bose and J.E. Hammerberg, J. Non-Cryst. Solids **61** & **62**, 1195 (1984); L.E. Ballentine and J.E. Hammerberg, Can. J. Phys. **62**, 692 (1984); S.K. Bose, L.E. Ballentine and J.E. Hammerberg, J. Phys. F **13**, 2089 (1983);
(b) L.E. Ballentine and M. Kolář, J. Phys. C: Solid State Phys. **19**, 981 (1986).

38. O.K. Andersen, G. Krier, R.W. Tank, C. Arcangeli, T. Dasgupta and O. Jepsen, this volume., see also O.K. Andersen,*Lectures on Methods of Electronic Structure Calculations*, Proceedings of the Miniworkshop on "Methods of Electronic Structure Calculations" and Working Group on "Disordered Alloys", ICTP, Trieste, Italy, (World Scientific Singapore, 1994), pp. 63-123.

MAGNETISM AND SPIN TUNNELING IN NANOSTRUCTURES

Alexander BRATKOVSKY
Hewlett-Packard Laboratories, 3500 Deer Creek Road, Palo Alto, CA 94304-1392,
alexb@hpl.hp.com

ABSTRACT

In the present paper different tunneling mechanisms in conventional and half-metallic ferromagnetic tunnel junctions are analyzed within the same general method. Theoretically calculated direct tunneling in iron group systems leads to about a 30% change in resistance, which is close but lower than experimentally observed values. It is shown that the larger observed values of the TMR might be a result of tunneling involving surface polarized states. We find that tunneling via resonant defect states in the barrier radically decreases the TMR (down to 4% with Fe-based electrodes), and a resonant tunnel diode structure would give a TMR of about 8%. With regards to inelastic tunneling, magnons and phonons exhibit opposite effects: one-magnon emission generally results in spin mixing and, consequently, reduces the TMR, whereas phonons are shown to enhance the TMR. The inclusion of both magnons and phonons reasonably explains an unusual bias dependence of the TMR.

The model presented here is applied qualitatively to half-metallics with 100% spin polarization, where one-magnon processes are suppressed and the change in resistance in the absence of spin-mixing on impurities may be arbitrarily large. Even in the case of imperfect magnetic configurations, the resistance change can be a few 1000 percent. Examples of half-metallic systems are CrO_2/TiO_2 and CrO_2/RuO_2, and an account of their peculiar band structures is presented. The implications and relation of these systems to CMR materials, which are nearly half-metallic, are discussed.

INTRODUCTION

Tunnel magnetoresistance (TMR) in ferromagnetic junctions, first observed more than a decade ago,[1,2] is of fundamental interest and potentially applicable to magnetic sensors and memory devices.[3] This became particularly relevant after it was found that the TMR for $3d$ magnetic electrodes reached large values at room temperature[4,5], and junctions demonstrated a non-volatile memory effect. These observations have ignited a world-wide effort towards using this effect in various applications, with memories and sensors being the most natural choices.

A simple model for spin tunneling has been formulated by Julliere[1] and further developed in Refs. [6,7]. This model is expected to work rather well for iron, cobalt, and nickel based metals, according to theoretical analysis[6] and experiments.[4] However, it disregards important points such as impurity-assisted and inelastic scattering, tunneling into surface states, and the reduced effective mass of carriers inside the barrier. These effects are important for proper understanding of the behavior of actual devices, like peculiarities in their $I - V$ curves, as considered in Ref. [8] and the present paper. I shall also discuss a couple of *half-metallic* systems which should in principle achieve the ultimate magnetoresistance at room temperatures and low fields.

ELASTIC AND INELASTIC TUNNELING, MODEL

The model that we will consider below includes a Hamiltonian for non-interacting conducting spin-split electrons $\mathcal{H}_0$, electron-phonon interaction $\mathcal{H}_{ep}$, and exchange interaction with localized d_l electrons $\mathcal{H}_x$, the later giving rise to the electron-magnon interaction. Impurities will be described by a short-range confining potential V_i,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{ep} + \mathcal{H}_x + \mathcal{H}_i, \tag{1}$$
$$\mathcal{H}_i = \sum_{\mathbf{n_i}} V_i(\mathbf{r} - \mathbf{n_i})$$

where $\mathbf{r}$ stands for the coordinate of the electron and $\mathbf{n_i}$ denotes the impurity sites.

The non-interacting part of the Hamiltonian $\mathcal{H}$ describes electrons in the ferromagnetic electrodes and insulating barrier according to the Schrödinger equation[7]

$$(\mathcal{H}_{00} - \mathbf{h} \cdot \hat{\sigma})\psi = E\psi, \tag{2}$$

where $\mathcal{H}_{00} = -(\hbar^2/2m_\alpha)\nabla^2 + U_\alpha$ is the single-particle Hamiltonian with $U(\mathbf{r})$ the potential energy, $\mathbf{h}(\mathbf{r})$ the exchange energy ($= 0$ inside the barrier), σ stands for the Pauli matrices; indices $\alpha=1$, 2, and 3 mark the quantities for left terminal, barrier, and right terminal, respectively ($\mathcal{H}_0$ is the expression in brackets). We shall also use the following notations to clearly distinguish between left and right terminal: $\mathbf{p} = \mathbf{k_1}$ and $\mathbf{k} = \mathbf{k_3}$. Solution to this problem in the limit of a thick barrier provides us with the basis functions for electrons in the terminals and barrier to be used in Bardeen's tunneling Hamiltonian approach.[9,10] We assume that all many-body interactions in the electrodes are included in the effective parameters of (2). To fully characterize tunneling we add to Bardeen's direct tunneling term $\mathcal{H}_T^0$ the contributions from $\mathcal{H}_x$ and $\mathcal{H}_{ep}$:

$$\mathcal{H}_T = \mathcal{H}_T^0 + \mathcal{H}_T^x + \mathcal{H}_T^{ep}, \tag{3}$$
$$\mathcal{H}_T^0 = \sum_{\mathbf{p},ka} T^0_{pa,ka} r^\dagger_{ka} l_{pa} + h.c.,$$
$$T^0_{pa,ka} = -\hbar^2/(2m_2) \int_\Sigma d\mathbf{A} \left(\bar{\psi}_{ka} \nabla\psi_{pa} - \nabla\bar{\psi}_{ka} \psi_{pa} \right); \tag{4}$$
$$\mathcal{H}_T^x = -\sum_{\alpha n,k,p} T^{J,\alpha}_{\mathbf{k},\mathbf{p}}(\mathbf{n}) \left[(S_n^3 - \langle S_n^3 \rangle)(r^\dagger_{k\uparrow} l_{p\uparrow} - r^\dagger_{k\downarrow} l_{p\downarrow}) + S_n^+ r^\dagger_{k\downarrow} l_{p\uparrow} + S_n^- r^\dagger_{k\uparrow} l_{p\downarrow} \right] + h.c.,$$
$$\mathcal{H}_T^{ep} = \sum_{\alpha a n,k,p} T^{ep,\alpha}_{\mathbf{k},\mathbf{p}}(\mathbf{q}) r^\dagger_{ka} l_{pa}(b_{q\alpha} - b^\dagger_{-q\alpha}) + h.c. \tag{5}$$

Here the surface Σ lies somewhere in the barrier and separates the electrodes, we have subtracted an average spin $S_n^3 - \langle S_n^3 \rangle$ in each of electrodes as part of the exchange potential, the exchange vertex is $T^J \sim J_\mathbf{n} \exp(-\kappa w)$, and the phonon vertex is related to the deformation potential D in the usual way $[T^{ep}(q) \sim \imath Dq(\hbar/2M\omega_q)^{1/2} \exp(-\kappa w)]$, where M is the atomic mass, $\mathbf{q}$ is the phonon momentum, $\mathbf{n}$ marks the lattice sites, and the vertices contain the square root of the barrier transparency.[10,11] The operators l_a and r_a annihilate electrons with spin a on the left and right electrodes, respectively. Two more things to note: (i) the summations over $\mathbf{p}$ and $\mathbf{k}$ always include the densities of initial g_{La} and final g_{Rb} states, that makes both exchange *and* phonon contributions spin-dependent, (ii) when the magnetic moments on the electrodes are at a mutual angle θ, one has to express the operator r w.r.t. the lab system and then use it in $\mathcal{H}_T$ (5).

118

The tunnel current will be calculated within the linear response formalism as[10]

$$I(V, t) = \frac{\imath e}{\hbar} \int_{-\infty}^{t} dt' \langle [dN_L(t)/dt, \mathcal{H}_T(t')] \rangle_0, \qquad (6)$$

where $N_L(t) = \sum_{\mathbf{pa}} l_{\mathbf{pa}}^{\dagger}(t) l_{\mathbf{pa}}(t)$ is the operator of the number of electrons on the left terminal in the interaction representation, $\langle\ \rangle_0$ stands for the average over $\mathcal{H}_0$,

$$\mathcal{H}_T(t) = \exp(-\imath e V t/\hbar) A(t) + h.c., \qquad A(t) = \sum_{\mathbf{pa,kb}} T_{\mathbf{pa,kb}}(t) r_{\mathbf{kb}}^{\dagger}(t) l_{\mathbf{pa}}(t),$$

the tunnel vertex T is derived for each term in (5), and V is the bias. We shall later consider impurity-assisted tunneling within the same general approach.

Elastic tunneling

We are now in position to calculate all contributions to the tunneling current, the simplest being direct elastic tunneling due to $\mathcal{H}_T^0$. It is worth noting that it can also be calculated from the transmission probabilities of electrons with spin a, $T_a = \sum_b T_{ab}$, which have a particularly simple form for a square barrier and *collinear* [parallel (P) or antiparallel (AP)] moments on the electrodes.[8] We obtain the following expression for the direct tunneling conductance, assuming $m_1 = m_3$ (below the effective mass in the barrier will be measured in units of m_1):

$$\frac{G^0}{A} = \frac{1}{A}\left(\frac{I}{V}\right)_{V\to 0} = G_{\mathrm{FBF}}^0 (1 + P_{\mathrm{FB}}^2 \cos(\theta)), \qquad (7)$$

$$G_{\mathrm{FBF}}^0 = \frac{e^2}{\pi\hbar}\frac{\kappa_0}{\pi w}\left[\frac{m_2\kappa_0(k_\uparrow + k_\downarrow)(\kappa_0^2 + m_2^2 k_\uparrow k_\downarrow)}{(\kappa_0^2 + m_2^2 k_\uparrow^2)(\kappa_0^2 + m_2^2 k_\downarrow^2)}\right]^2 e^{-2\kappa_0 w}, \quad \text{and} \qquad (8)$$

$$P_{\mathrm{FB}} = \frac{k_\uparrow - k_\downarrow}{k_\uparrow + k_\downarrow}\frac{\kappa_0^2 - m_2^2 k_\uparrow k_\downarrow}{\kappa_0^2 + m_2^2 k_\uparrow k_\downarrow}, \qquad (9)$$

where P_{FB} is the effective polarization of the ferromagnetic (F) electrode in the presence of the barrier (B), $\kappa_0 = [2m_2(U_0 - E)/\hbar^2]^{1/2}$, and U_0 is the top of the barrier. Eq. (7) corrects an expression derived earlier[7] for the effective mass of the carriers in the barrier. By taking a typical value of G/A =4-5 $\Omega^{-1}\mathrm{cm}^{-2}$ (Ref. [4]), $k_\uparrow = 1.09\text{Å}^{-1}$, $k_\downarrow = 0.42\text{Å}^{-1}$, $m_1 \approx 1$ (for itinerant d electrons in Fe)[6] and a typical barrier height for Al_2O_3 (measured from the Fermi level μ) $\phi = U_0 - \mu = 3\mathrm{eV}$, and the thickness $w \approx 20$ Å, one arrives at the following estimate for the effective mass in the barrier: $m_2 \approx 0.4$.[13] These values give the renormalized polarization $P_{\mathrm{FeB}} = 0.28$, which is less than the bulk value for iron $P_{\mathrm{Fe}} = 0.4$ (Ref. [3,4]). Note that the neglect[7] of the mass correction makes $P_{\mathrm{FeB}} < 0$, a result which is not corroborated by experimental evidence where the polarization in all systems studied was found positive, $P > 0$. The majority spin electrons in all cases were predominant in the tunnel current (Ref. [3], p.204).

In the standard approximation of a rectangular shape the barrier height is $U_0 = \frac{1}{2}(\phi_L + \phi_R - eV)$ and this leads to a quick rise of the conductance with bias, $G^0(V) = G^0 + const \cdot V^2$ at small V (ϕ_L and ϕ_R are the work functions of the electrodes). In practice, the barrier parameters should be extracted from independent experiments, such as internal photoemission, etc., but here we are concerned with the generic behavior, where the present formalism is sufficient for qualitative and even semi-quantitative analysis. Since the barrier shape depends in a non-trivial manner on image forces, the calculations have been performed numerically with the actual barrier shape at finite temperatures (Fig. 1).
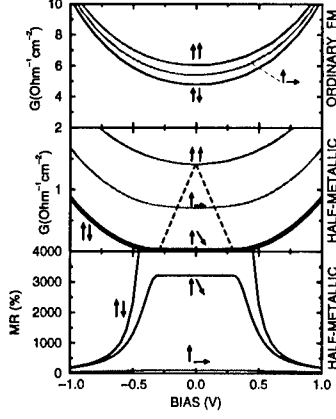
Figure 1: Conductance and magnetoresistance of ferromagnetic tunnel junctions versus bias. Top panel: conventional (Fe-based) tunnel junction (for parameters see text). Middle panel: half-metallic electrodes. Bottom panel: magnetoresistance for the half-metallic electrodes. Dashed lines show schematically a region where a half-metallic gap in the minority spin states is controlling the transport. Even for imperfect antiparallel alignment ($\theta = 160°$, marked $\uparrow\searrow$), the magnetoresistance for half-metallics (bottom panel) exceeds 3000% at biases below the threshold V_c. All calculations have been performed at 300K with the inclusion of multiple image potential and exact transmission coefficients. Parameters are described in the text.

We note that the (undesirable) downward renormalization of the polarization rapidly goes with diminishing effective carrier mass in the barrier. The renormalization is completely absent in half-metallic ferromagnets with $\mathrm{Re}k_\downarrow = 0$, as we shall discuss below.

We define the magnetoresistance as the relative change in contact conductance with respect to the change of mutual orientation of spins from parallel (G^P for $\theta = 0$) to antiparallel (G^{AP} for $\theta = 180°$) as

$$MR = (G^P - G^{AP})/G^{AP} = 2P_{FB}P'_{FB}/(1 - P_{FB}P'_{FB}). \tag{10}$$

The most striking feature of Eqs. (3),(4) is that the MR tends to infinity for vanishing $\mathrm{Re}k_\downarrow$, i.e. when both electrodes are made of a 100% spin-polarized material ($P = P' = 1$), because of a gap in the density of states (DOS) for minority carriers up to their conduction band minimum $E_{CB\downarrow}$. Then G^{AP} vanishes together with the tunnel probability, since there is a zero DOS at $E = \mu$ for both spin directions.

Such half-metallic behavior is rare, but some materials possess this amazing property, most interestingly the oxides CrO_2 and Fe_3O_4.[14] These oxides have potential for future applications in combination with lattice-matching materials, as we shall illustrate below.

A more accurate analysis of the $I - V$ curve requires a numerical evaluation of the tunnel current for arbitrary biases and image forces, and the results are shown in Fig. 1. The top panel in Fig. 1 shows $I - V$ curves for an iron-based F-B-F junction with the above-mentioned parameters. The value of TMR is about 30% at low biases and steadily decreases with increased bias. In a half-metallic case ($\mathrm{Re}k_\downarrow = 0$, Fig. 1, middle panel, where a threshold $eV_c = E_{CB\downarrow} - \mu = 0.3$ eV has been assumed), we obtain *zero* conductance G^{AP} in the AP configuration at biases lower than V_c. It is easy to see that above this threshold $G^{AP} \propto (V - V_c)^{5/2}$ at temperatures much smaller than eV_c.[8] Thus, for $|V| < V_c$ in the

120
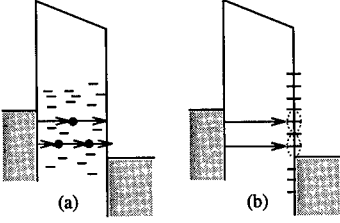
Figure 2: Schematic of tunneling via chains of the localized states in the barrier (a) and into the localized surface states (b).

AP geometry one has $MR = \infty$. In practice, there are several effects that reduce this MR to some finite value, notably an imperfect AP alignment of moments in the electrodes. However, from the middle and the bottom panels in Fig. 1 we see that even at 20° deviation from the AP configuration, the value of MR exceeds 3,000% within the half-metallic gap $|V| < V_c$, and this is indeed a very large value.

Impurity-assisted tunneling

An important aspect of spin-tunneling is the effect of tunneling through the defect states in the (amorphous) oxide barrier (Fig. 2). Since the devices under consideration are very thin, their $I - V$ curves and MR should be very sensitive to defect resonant states in the barrier with energies close to the chemical potential, forming "channels" with the nearly periodic positions of impurities (Fig. 2).[15] Generally, channels with one impurity (most likely to dominate in thin barriers) would result in a monotonous behavior of the $I - V$ curve, whereas channels with *two or more* impurities would produce intervals with negative differential conductance.[15]

Impurity-assisted spin tunneling at zero temperature (at non-zero T one should include an integration with the Fermi functions) has a resonant form[15,8]

$$G_a = \frac{2e^2}{\pi\hbar} \sum_i \frac{\Gamma_{La}\Gamma_{Ra}}{(E_i - \mu)^2 + \Gamma^2}, \tag{11}$$

where $\Gamma = \Gamma_{La} + \Gamma_{Ra}$ is the total width of the resonance given by the sum of the partial widths Γ_L (Γ_R) corresponding to electron tunneling from the impurity state at the energy E_i to the left (right) terminal. For the tunnel width we have

$$\Gamma_{(L,R)a} = 2\pi^2\kappa_0(\hbar^2/m_2)^2 \sum_{\mathbf{k}_{(L,R)a}} |\psi_{\mathbf{k}_{(L,R)a}}(\mathbf{n}_i)|^2 \delta(E_\mathbf{k} - E_i), \tag{12}$$

where $\psi_{\mathbf{k}_{(L,R)a}}(\mathbf{n}_i)$ is the value of the electrode wave function, exponentially decaying into the barrier, at an impurity site $\mathbf{n}_i$. For a rectangular barrier we have[8]

$$\Gamma_{La} = \epsilon_i \frac{2m_2 k_a}{\kappa_0^2 + m_2^2 k_a^2} \frac{e^{-\kappa_0(w+2z_i)}}{\kappa_0(\frac{1}{2}w + z_i)}, \tag{13}$$

where z_i is the coordinate of the impurity with respect to the center of the barrier, $\epsilon_i = \hbar^2\kappa_0^2/(2m_2)$. For e.g. P configuration and electrodes of the same material, the conductance would then be proportional to $\left[(E_i - \mu)^2 + 4\Gamma_{0a}^2 \cosh^2(2\kappa_0 z_i)\right]^{-1}$, where Γ_{0a} equals

(13) without the factor $\exp(-2\kappa_0 z_i)$ [c.f. Eq. (15)]. The conductance has a sharp maximum $(= e^2/(2\pi\hbar))$ when $\mu = E_i$ and $\Gamma_L = \Gamma_R$, i.e. for the symmetric position of the impurity in the barrier $|z_i| < 1/\kappa_0$ in a narrow interval of energies $|\mu - E_i| < \Gamma$. Averaging over energies and positions of impurities in Eq. (11), and considering a general configuration of the magnetic moments on the terminals, we get the following formula for impurity-assisted conductance in the leading order in $\exp(-\kappa w)$:

$$\frac{G^1}{A} = G_{\text{imp}}^1 (1 + \Pi_{\text{FB}}^2 \cos(\theta)), \qquad (14)$$

where we have introduced the quantities

$$G_{\text{imp}}^1 = \frac{e^2}{\pi\hbar} N_1, \quad N_1 = \pi^2 \nu \Gamma_1/\kappa_0,$$

$$\Gamma_1 = \epsilon_i \frac{e^{-\kappa_0 w}}{\kappa_0 w} (r_\uparrow + r_\downarrow)^2, \quad \Pi_{\text{FB}} = (r_\uparrow - r_\downarrow)/(r_\uparrow + r_\downarrow), \quad \text{and}$$

$$r_a = [m_2 \kappa_0 k_a/(\kappa_0^2 + m_2^2 k_a^2)]^{1/2}, \qquad (15)$$

with N_1 being the effective number of one-impurity channels per unit area, and Π_{FB} is the 'polarization' of the impurity channels. When the total number of one-impurity channels $\mathcal{N}_1 = N_1 A \gg 1$, then the conductance will be a self-averaged quantity, otherwise it will depend on a specific arrangement of impurities (regime of mesoscopic fluctuations).

Comparing the direct and the impurity-assisted contributions to conductance, we see that the latter dominates when the impurity density of states $\nu \geq (\kappa_0/\pi)^3 \epsilon_i^{-1} \exp(-\kappa_0 w)$, and in our example a crossover takes place at $\nu \geq 10^{17} \text{cm}^{-3}\text{eV}^{-1}$. When the resonant transmission dominates, the magnetoresistance is given by

$$MR_1 = 2\Pi \, \Pi'/(1 - \Pi \, \Pi'), \qquad (16)$$

which is just 4% in the case of Fe. Thus, we have a drastic reduction of the TMR due to non-magnetic impurities in the tunnel barrier, and in the case of magnetic impurities the TMR will be even smaller.

With standard ferromagnetic electrodes, the conductance is exponentially enhanced $[G^1 \propto \exp(-\kappa_0 w)$, whereas $G^0 \propto \exp(-2\kappa_0 w)]$ but the magnetoresistance is reduced in comparison with the 'clean' case of a low concentration of defect levels. These predictions[8] have been confirmed by recent experiments.[12,16]

With further increase of the defect density and/or the barrier width, the channels with two- and more impurities will become more effective than one-impurity channels described above, as has been known for quite a while.[17,15] The contribution of the many-impurity channels, generally, will result in the appearance of irregular intervals with negative differential conductance on the $I - V$ curve.[15] Thus, the two-impurity channels define random fluctuations of current with bias. This is due to the fact that the energy of defect states depends on bias as $\epsilon_i = \epsilon_i^0 + eV z_i/w$. With increasing bias (i) the total number of two-impurity channels increases but (ii) some of these channel go off resonance and reduce their conductance. Accidentally, the number of two-impurity channels going off resonance may become larger than a number of new channels, leading to a suppressed overall conductance. If we denote by Γ_2 the width of the two-impurity channels, then the fluctuations would obviously occur on a scale $\Delta V < \Gamma_2/e$. Then, according to standard arguments, the change in current will be

$$\frac{\Delta I}{I} = \frac{\Delta V}{V} \pm \left(\frac{e\Delta V}{\Gamma_2}\right) \mathcal{N}^{-1/2}, \qquad (17)$$

where $\mathcal{N} = eV\mathcal{N}_2/\Gamma_2$ is the number of the two-impurity channels contributing at the bias $V > \Gamma_2/e$, $\mathcal{N}_2$ is the total number of the two-impurity channels, $\mathcal{N}_2 = A\pi^3 w^3 \nu^2 \Gamma_2^2 \kappa_0^{-1}$, and $\Gamma_2 = (4\epsilon_i \Gamma_1/(\kappa_0 w))^{1/2}$.[15] When $eV/\Gamma_2 > \mathcal{N}_2(\kappa_0 w)^2$, then the second (random) term in (17) exceeds the first term, and this leads to random intervals with negative differential conductance.

Obviously, with increasing temperature or/and bias in thick enough barriers longer and longer impurity channels will be 'turned on'. A corresponding microscopic model should include impurity states coupled to a phonon bath, and such a model has been solved in Ref. [18]. The authors found an average conductance due to n-impurity chain in the limit $eV \ll T$:

$$\frac{G_n(T)}{G_1} \sim (\nu \kappa_0^{-2} wT)^{n-1} (\lambda T/\epsilon_0)^{(n-1)/(n+1)} \exp\left(\frac{n-1}{n+1} \kappa_0 w\right), \tag{18}$$

where the dimensionless parameter $\lambda = D^2 \epsilon_0^2/(\hbar^3 \rho s^5)$, ρ is the mass density, s is the velocity of sound. As follows from (18) there is a sequence of specific power law dependencies of G on temperature. It is also evident that the hopping exponent strongly favors many-impurity chains. For $n = 2$ we have $G_2(T) \propto T^{4/3}$, then $G_3(T) \propto T^{5/2}$, and so on. In the opposite limiting case $eV \gg T$ the result is:[18] $G_2(V) \propto V^{4/3}$, and this crossover behavior is indeed in very good agreement with experiment on a-Si barriers.[19]

One may try to fabricate a resonant tunnel diode (RTD) structure to sharply increase the conductance of a system. We can imagine an RTD structure with an extra thin non-magnetic layer placed between two oxide barrier layers producing a resonant level at some energy E_r. The only difference from the previous discussion is the effectively $1D$ character of the transport in the RTD in comparison with $3D$ impurity-assisted transport. However, the transmittance will have the same resonant form as in (11). The estimated magnetoresistance in the RTD geometry is, with the use of (11),

$$MR_{\mathrm{RTD}} = \left[(r_\uparrow^2 - r_\downarrow^2)/(2r_\uparrow r_\downarrow)\right]^2, \tag{19}$$

which is about 8% for Fe electrodes. We see that the presence of random impurity levels or a single resonant level reduces the value of the magnetoresistance as compared with direct tunneling.

The general reason for the MR being reduced even by non-magnetic impurities is the downward renormalization of spin polarization of tunneling current by *non*-magnetic insulator (proximity effect). Since the exchange potential vanishes in non-magnetic insulator like Al_2O_3, matching of the corresponding wave functions results in a reduced difference in the density of majority and minority states, i.e. reduced polarization. Same, of course, is true of electron states in electrodes overlapping with a non-polarized defect state in the barrier. As we have seen, this proximity effect is enough to reduce MR down to small values, and if we were to include exchange effects (spin mixing) on impurities the MR will be even less. This prediction[8] has indeed been confirmed experimentally.[12,16]

Roughness

As we have seen, the conductance is dominated by the exponentially small barrier transparency, $\propto \exp(-2w(\kappa_0^2 + \mathbf{k}_\parallel^2)^{1/2})$, so that the contribution comes mainly from electrons tunneling perpendicular to the barrier, i.e. with small parallel momenta $|\mathbf{k}_\parallel| < (\kappa_0/w)^{1/2}$. For barriers with a rough interface $w = \overline{w} + h$, where h is the height of asperities and $\overline{w}$

is the average barrier thickness. Each asperity will contribute a factor of $\exp(2\kappa_0 h)$ to the conductance, which we have to average. We assume a normal distribution for roughness, $P(h) = (2\pi h_0^2)^{-1/2} \exp\left(-h^2/(2h_0^2)\right)$. Then, the average conductance $\overline{G}$ becomes

$$G = \overline{G} \int_{-\infty}^{\infty} dh \exp(2\kappa_0 h) P(h) = \overline{G} \exp(2\kappa_0^2 h_0^2) \propto \exp[-2\kappa_0(\overline{w} - \kappa h_0^2)]. \tag{20}$$

This result means that the effective thickness of the barrier is reduced by κh_0^2 in comparison with the observed average thickness $\overline{w}$. The generalization for the case of correlated roughness is straightforward and does not change this result.

Tunneling via Surface States

Direct tunneling, as we have seen, gives TMR of about 30%, whereas in recent experiments TMR is well above this value, approaching 40%.[12,23] As we shall see shortly, this moderate difference is unlikely to come from the inelastic processes. Up to now we have disregarded the possibility of localized states at metal-oxide interfaces (Fig. 2). Keeping in mind that the usual barrier AlO_x is amorphous, the density of such states may well exceed that at typical semiconductor-oxide surfaces. If this is true, then we have to take into account tunneling into/from those states. If we assume that electrons at the surface are confined by a short-range potential then we can estimate the tunneling matrix elements as described above. The corresponding tunneling MR is given by[11]

$$
\begin{aligned}
\frac{G_{\mathrm{bs}}(\theta)}{A} &= \frac{e^2}{\pi\hbar} B\overline{D}_s(1 + P_{\mathrm{FB}}P_s\cos(\theta)), \\
P_s &= \frac{D_{s\uparrow} - D_{s\downarrow}}{D_{s\uparrow} + D_{s\downarrow}}, \\
\overline{D}_s &= \frac{1}{2}(D_{s\uparrow} + D_{s\downarrow}), \\
B &= \frac{4\pi\epsilon_s}{\kappa_0 w} \frac{m_2\kappa_0(k_\uparrow + k_\downarrow)(\kappa_0^2 + m_2^2 k_\uparrow k_\downarrow)}{(\kappa_0^2 + m_2^2 k_\uparrow^2)(\kappa_0^2 + m_2^2 k_\downarrow^2)} \exp(-2\kappa_0 w), \tag{21}
\end{aligned}
$$

where P_s is the polarization and $\overline{D}_s$ is the average density of surface states, $\epsilon_s = \hbar^2\kappa_0^2/(2m_2)$. The corresponding magnetoresistance would be $MR_{\mathrm{bs}} = 2P_{\mathrm{FB}}P_s/(1 - P_{\mathrm{FB}}P_s)$.

Comparing (21) with (7), we see that the bulk-to-surface conductance exceeds bulk-to-bulk tunneling at moderate densities of surface states $D_s > D_{sc} \sim 10^{13}\mathrm{cm}^{-2}\mathrm{eV}^{-1}$ per spin, comparable to those found in MOSFET structures.

If on both sides of the barrier the density of surface states is above critical value D_{sc}, the magnetoresistance will be due to surface-to-surface tunneling with a value given by

$$MR_{\mathrm{ss}} = 2P_{s1}P_{s2}/(1 - P_{s1}P_{s2}),$$

and if the polarization of surface states is larger than in the bulk, as is often the case even for imperfect surfaces,[20] then it would result in enhanced TMR. This mechanism may be even more relevant for Fe/Si and other ferromagnet-semiconductor structures.[21]
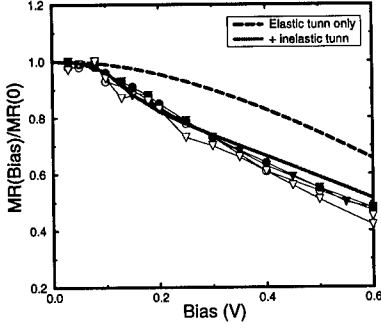
Figure 3: Fit to experimental data for the magnetoresistance of $Co/Al_2O_3/NiFe$ tunnel junctions [12] with inclusion of elastic and inelastic (magnons and phonons) tunneling. The fit gives for magnon DOS $\propto \omega^{0.65}$ which is close to the standard spectrum $\propto \omega^{1/2}$.

INELASTIC TUNNELING, 'ZERO-BIAS' ANOMALY

So far we have disregarded all inelastic processes, such as phonon emission by the tunneling electrons. These processes were long thought to be responsible for a so-called 'zero-bias' anomaly observed in a variety of non-magnetic[22] and magnetic junctions.[12,23] Magnetism in electrodes introduces new peculiarities into the problem, which we will now discuss. The obvious one is related to emission of magnons. At temperatures well below the Curie temperature and not very large biases, one can describe spin excitations by introducing magnons. Then the calculations of exchange- and phonon-assisted currents become very similar. Thus, we obtain from (6) and (5) the following expression for magnon-assisted current in e.g. parallel configuration (corresponding expressions can be easily found for other configurations as well):

$$
\begin{aligned}
I_P^x(V,T) &= \frac{2\pi e}{\hbar} \sum_{q\alpha} X^\alpha \bigg(g_\uparrow^L g_\downarrow^R \ (eV + \omega) \left[\frac{N_\omega}{1 - \exp(-\beta(eV + \omega))} + \frac{N_\omega + 1}{1 - \exp(\beta(eV + \omega))} \right] \\
&+ g_\downarrow^L g_\uparrow^R \ (eV - \omega) \left[\frac{N_\omega + 1}{1 - \exp(-\beta(eV - \omega))} + \frac{N_\omega}{1 - \exp(\beta(eV - \omega))} \right] \bigg),
\end{aligned}
\tag{22}
$$

where $N_\omega = [\exp(\beta\omega) - 1]^{-1}$, $\beta = 1/T$ is the inverse temperature, $\omega = \omega_q^\alpha$ and X^α is the magnon incoherent vertex related to the $|T_{p,k}^{x,\alpha}(2S_n/N)^{1/2}|^2$ (5) with all momenta parallel to the barrier integrated out.[11] To get this expression, we have also assumed that the electron densities of states g in (22) vary on a larger scale than the bosonic contributions do, and, therefore, substituted them by representative values at the nominal Fermi levels. If there are some fine features in the electron DOS, then the integral over electron energies should remain, thus necessarily smoothing out any such fine features in the electron DOS.

For the limiting case of $T = 0$, we obtain for magnon-assisted inelastic tunneling current:

$$
\begin{aligned}
I_P^x &= \frac{2\pi e}{\hbar} \sum_\alpha X^\alpha g_\downarrow^L g_\uparrow^R \int d\omega \rho_\alpha^{mag}(\omega)(eV - \omega)\theta(eV - \omega), \\
I_{AP}^x &= \frac{2\pi e}{\hbar} \left[X^R g_\uparrow^L g_\uparrow^R \int d\omega \rho_R^{mag}(\omega)(eV - \omega)\theta(eV - \omega) \right.
\end{aligned}
$$

$$+ X^L g_\downarrow^L g_\downarrow^R \int d\omega \rho_L^{mag}(\omega)(eV - \omega)\theta(eV - \omega)\Big]. \tag{23}$$

where $\theta(x)$ is the step function, $\rho_\alpha^{mag}(\omega)$ is the magnon density of states that has a general form $\rho_\alpha^{mag}(\omega) = (\nu + 1)\omega^\nu/\omega_0^{\nu+1}$, ν can be used as a fitting parameter to define a dispersion of the relevant magnons, and ω_0 is the maximum magnon frequency.

For phonon-assisted current at $T = 0$ we have

$$I_P^{ph} = \frac{2\pi e}{\hbar} \sum_{a\alpha} g_a^L g_a^R \int d\omega \rho_\alpha^{ph}(\omega) P^\alpha(\omega)(eV - \omega)\theta(eV - \omega), \tag{24}$$

$$I_{AP}^{ph} = \frac{2\pi e}{\hbar} \sum_{a\alpha} g_a^L g_{-a}^R \int d\omega \rho_\alpha^{ph}(\omega) P^\alpha(\omega)(eV - \omega)\theta(eV - \omega). \tag{25}$$

One can show that the ratio of phonon to exchange vertex is $P(\omega)/X = \gamma\omega/\omega_D$, where γ is a constant depending on the ratio between deformation potential and exchange constants,[11] and ω_D is the Debye frequency.

The elastic and inelastic contributions together will define the total junction conductance $G = G(V, T)$ as a function of the bias V and temperature T. We find that the inelastic contributions from magnons and phonons (23)-(25) grow as $G^x(V, 0) \propto (|eV|/\omega_0)^{\nu+1}$ and $G^{ph}(V, 0) \propto (|eV|/\omega_D)^4$ at low biases. These contributions saturate at higher biases: $G^x(V, 0) \propto 1 - \frac{\nu+1}{\nu+2}\frac{\omega_0}{|eV|}$ at $|eV| > \omega_0$; $G^{ph}(V, 0) \propto 1 - \frac{4}{5}\frac{\omega_D}{|eV|}$ at $|eV| > \omega_D$. This behavior would lead to sharp features in the $I - V$ curves on a scale of 30-100 mV (Fig. 3).

It is important to highlight the opposite effects of phonons and magnons on the TMR. If we take the case of the same electrode materials and denote $D = g_\uparrow$ and $d = g_\downarrow$ then we see that $G_P^x(V, 0) - G_{AP}^x(V, 0) \propto -(D - d)^2(|eV|/\omega_0)^{\nu+1} < 0$, whereas $G_P^{ph}(V, 0) - G_{AP}^{ph}(V, 0) \propto +(D - d)^2(|eV|/\omega_D)^4 > 0$, i.e. spin-mixing due to magnons *kills*, whereas the phonons tend to *enhance* the TMR.[24]

Finite temperature gives contributions of the same respective sign as written above. For magnons: $G_P^x(0, T) - G_{AP}^x(0, T) \propto -(D - d)^2(-TdM/dT) < 0$, where $M = M(T)$ is the magnetic moment of the electrode at a given temperature T. The phonon contribution is given by a standard Debye integral with the following results: $G_P^{ph}(0, T) - G_{AP}^{ph}(0, T) \propto +(D - d)^2(T/\omega_D)^4 > 0$ at $T \ll \omega_D$, and a linear temperature dependence at high temperatures $G_P^{ph}(0, T) - G_{AP}^{ph}(0, T) \propto +(D - d)^2(T/\omega_D)$ at $T \gg \omega_D$.[11] We note again the opposite effect of magnons and phonons on the tunneling magnetoresistance.

It is worth mentioning that the magnon excitations are usually cut off by e.g. the anisotropy energy K_{an} at $\omega_c = 2g\mu_B K_{an}/M_s$, where M_s is the saturation magnetization. Therefore, at low temperatures $T \ll \omega_c$ the magnon finite temperature contribution to conductance will be exponentially small, $\propto \exp(-\omega_c/T)$, and the conductance at zero bias will be almost independent on temperature at $T \ll (\omega_c, \omega_D)$.

We have not included Kondo[25] and other correlation effects that might contribute at very low biases, since they usually do not help to quantitatively fit the data.[19]

The role of phonons is illustrated by my fit to recent experiments carried out at HPL:[12] it appears that only after including phonons is it possible to get a sensible fit to the magnon DOS with $\nu = 0.65$, which is close to the bulk value $\frac{1}{2}$ and $\gamma \approx 0.1$ (Fig. 3).

100% POLARIZATION

It is very important that *in the case of half-metallics* $r_\downarrow = 0$, $\Pi_{FB} = 1$, and even with an imperfect barrier magnetoresistance can, at least in principle, reach any value limited
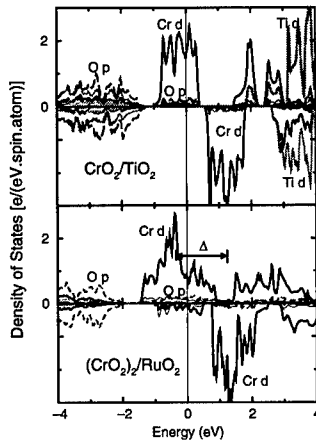
Figure 4: Density of states of CrO_2/TiO_2 (top panel) and $(CrO_2)_2/RuO_2$ (bottom panel) half-metallic layered structures calculated with the use of the LMTO method.

by only spin-flip processes in the barrier/interface and/or misalignment of moments in the half-metallic ferromagnetic electrodes.[8] We should note that the *one-magnon* excitations in half-metallics are suppressed by the half-metallic gap, as immediately follows from our discussion in the previous section. Spin-mixing can only occur on magnetic impurities in the barrier or interface, because the allowed *two*-magnon excitations in the electrodes do not result in spin-mixing.

Therefore, these materials should combine the best of both worlds: very large magnetoresistance with enhanced conductance in tunnel MR junctions. One should be aware, however, that defects in the barrier (like unpaired electrons) will induce spin flips, so the magnetoresistance could vanish with an increasing concentration of defects. In the case of conventional systems (e.g. NiFe electrodes) we have seen, however, that resonant tunneling significantly reduces the tunnel MR by itself, so the possibility of improving the conductance and still having a very large magnetoresistance resides primarily with half-metallics. I shall finish with a couple of examples of systems with half-metallic behavior, CrO_2/TiO_2 and CrO_2/RuO_2[8] (Fig. 4). They are based on half-metallic CrO_2, and all materials have the rutile structure with almost perfect lattice matching, which should yield a good interface and should help in keeping the system at the desired stoichiometry. TiO_2 and RuO_2 are used as the barrier/spacer oxides. The half-metallic behavior of the corresponding multilayer systems is demonstrated by the band structures calculated within the linear muffin-tin orbitals method (LMTO) in a supercell geometry with [001] growth direction and periodic boundary conditions. The calculations show that CrO_2/TiO_2 is a perfect half-metallic, whereas $(CrO_2)_2/RuO_2$ is a weak half-metallic, since there is some small DOS around E_F, and an exact gap opens up at about 0.58 eV above the Fermi level (Fig. 4). In comparison, there are only states in the majority spin band at the Fermi level in CrO_2/TiO_2. An immediate consequence of the fact that minority spin bands are fully occupied is an exact *integer* value of the magnetic moment in the unit cell ($=2\mu_B$/Cr in CrO_2/TiO_2), and this property is a simple check for possible *new* half-metallics.

The electronic structure of CrO_2/TiO_2 is very interesting in that it has a half-metallic

gap which is 2.6 eV wide and extends on both sides of the Fermi level, where there is a gap either in the minority *or* majority spin band. Thus, an huge magnetoresistance should in principle be seen not only for electrons at the Fermi level biased up to 0.5 eV, but also for *hot* electrons starting at about 0.5 eV above the Fermi level. We note that states at the Fermi level are a mixture of $Cr(d)$ and $O(2p)$ states, so that $p - d$ interaction within the first coordination shell produces a strong hybridization gap, and the Stoner spin-splitting moves the Fermi level right into the gap for minority carriers (Fig. 4).

An important difference between the two spacer oxides is that TiO_2 is an insulator whereas RuO_2 is a good metallic conductor. Thus, the former system can be used in a tunnel junction, whereas the latter will form a metallic multilayer. In the latter case the physics of conduction is different from tunneling but the effect of vanishing phase volume for transmitted states still works when current is passed through such a system *perpendicular to planes*. For the P orientation of moments on the electrodes, CrO_2/RuO_2 would have a normal metallic conduction, whereas in the AP one we expect it to have a semiconducting type of transport, with a crossover between the two regimes. One interesting possibility is to form three-terminal devices with these systems, like a spin-valve transistor,[26] and check the effect in the hot-electron region. CrO_2/TiO_2 seems to a be a natural candidate to check the present predictions about half-metallic behavior and for a possible large tunnel magnetoresistance. An important advantage of these systems is almost perfect lattice matching at the oxide interfaces. The absence of such a match of the conventional Al_2O_3 barrier with Heusler half-metallics (NiMnSb and PtMnSb) may have been among other reasons for their moderate performance.[27]

By using all-oxide half-metallic systems, as described herein, one may bypass many materials issues. Then, the main concerns for achieving a very large value of magnetoresistance will be spin-flip centers and imperfect alignment of moments. As for conventional tunnel junctions, the present results show that the presence of defect states in the barrier, or a resonant state like in a resonant tunnel diode type of structure, reduces their magnetoresistance by several times but may dramatically increase the current through the structure.

Finally, we can mention the CMR materials. Experiment[28] and LDA calculations[29] indicate that manganites are close to half-metallic behavior as a result of a significant spin-splitting presumably due to significant Hund's rule coupling on Mn. Manganites are strongly correlated materials, likely with electronic phase separation,[30] which makes their study a real challenge. There are a number of studies of systems, where transport is going across grain boundaries or between MnO_2 layers in tailored derivatives of the perovskite phase.[31] A hope is that some of these structures with manganites might operate at low fields and reasonably high temperatures.[32] The low field (below 1000 Oe) TMR in polycrystalline $La_{2/3}Sr_{1/3}MnO_3$ perovskite and $Tl_2Mn_2O_7$ pyrochlore is about 30% and is likely due to intergrain carrier transport. It would be interesting to apply the results of the present work to tunneling phenomena in the CMR-based layered/inhomogeneous structures. For instance, CrO_2 junctions would help to check on the relevance of the half-metallic behavior to conduction in the CMR materials. In particular, it should be signaled by a plateau in the tunneling magnetoresistance as a function of bias within the half-metallic band gap (Fig. 1).

REFERENCES

1. M. Julliere, Phys. Lett. **54A**, 225 (1975).
2. S. Maekawa and U. Gäfvert, IEEE Trans. Magn. **18**, 707 (1982).
3. R. Meservey and P.M. Tedrow, Phys. Reports **238**, 173 (1994).
4. J.S. Moodera *et al.*, Phys. Rev. Lett. **74**, 3273 (1995); J. Appl. Phys. **79**, 4724 (1996).
5. T. Miyazaki and N. Tezuka, J. Magn. Magn. Mater. **139**, L231 (1995).
6. M.B. Stearns, J. Magn. Magn. Mater. **5**, 167 (1977); Phys. Rev. B **8**, 4383 (1973).
7. J.C. Slonczewski, Phys. Rev. B **39**, 6995 (1989).
8. A.M. Bratkovsky, Phys. Rev. B **56**, 2344 (1997); JETP Lett. **65**, 452 (1997).
9. J. Bardeen, Phys. Rev. Lett., **6**, 57 (1961).
10. G.D. Mahan, Many-Particle Physics, 2nd ed., Plenum, New York, 1990, Ch. 9; C.B. Duke, Tunneling in Solids, Academic Press, New York, 1969, Ch. 7.
11. A.M. Bratkovsky (to be published).
12. J.Nickel, T.Anthony, and J. Brug, private communication.
13. Q.Q. Shu and W.G. Ma, Appl. Phys. Lett. **61**, 2542 (1992) give even smaller $m_2 = 0.2$ for $Al-Al_2O_3$-metal junctions.
14. V.Y. Irkhin and M.I. Katsnelson, Sov. Phys. - Uspekhi **164**, 705 (1994).
15. A.I. Larkin and K.A. Matveev, Zh. Eksp. Teor. Fiz. **93**, 1030 (1987); I.M. Lifschitz and V.Ya. Kirpichenkov, Zh. Eksp. Teor. Fiz. **77**, 989 (1979).
16. R. Jansen and J.S. Moodera (1997), to be published.
17. M. Pollak and J.J. Hauser, Phys. Rev. Lett. **31**, 1304 (1973); A.V. Tartakovskii *et al.*, Sov. Phys. Semicond. **21**, 370 (1987); E.I. Levin *et al.*, Sov. Phys. Semicond. **22**, 401 (1988); J.B. Pendry, J. Phys. C **20**, 733 (1987).
18. L.I. Glazman and K.A. Matveev, Sov. Phys. JETP **67**, 1276 (1988).
19. Y. Xu, D. Ephron, and M. Beasley, Phys. Rev. B **52**, 2843 (1995).
20. A.V. Smirnov and A.M. Bratkovsky, Phys. Rev. B **54**, R17371 (1996); *ibid.*, **55**, 14434 (1997).
21. A. Chaiken, R.P. Michel, and M.A. Wall, Phys. Rev. B **53**, 5518 (1996).
22. C.B. Duke, S.D. Silverstein, and A.J.Bennett, Phys. Rev. Lett. **19**, 315 (1967).
23. J. Moodera, private communication.
24. In a recent attempt to explain the $I-V$ curves of ferromagnetic junctions S. Zhang, P.M. Levy, A.C. Marley, and S.S.P. Parkin [Phys. Rev. Lett. **79**, 3744 (1997)] have apparently neglected a strong bias dependence of the direct tunneling and did not consider the effect of phonons.
25. J. Appelbaum, Phys. Rev. Lett. **17**, 91 (1966).
26. D.J. Monsma *et al.*, Phys. Rev. Lett. **74**, 5260 (1995).
27. C.T. Tanaka and J.S. Moodera, J. Appl. Phys. **79**, 6265 (1996).
28. Y. Okimoto *et al.*, Phys. Rev. B **55**, 4206 (1997); Phys. Rev. Lett. **75**, 109 (1995).
29. W.E. Pickett and D.J. Singh, Phys. Rev. B **53**, 1146 (1996); D.J. Singh, Phys. Rev. B **5**, 313 (1997).
30. G. Allodi *et al.* Phys. Rev. B **56**, 6036 (1997); E.L. Nagaev, JETP Lett. **6**, 484 (1967); Phys. Rev. B **54**, 16608 (1996); *ibid.* **56**, 14583 (1997).
31. M.K. Gubkin *et al.* Phys. Sol. State **35**, 728 (1993); H.Y. Hwang *et al.*, Phys. Rev. Lett. **77**, 2041 (1996); J.Z. Sun *et al.*, Appl. Phys. Lett. **69**, 3266 (1996); T. Kimura *et al.*, Science **274**, 1698 (1996).
32. H.Y. Hwang and S.W. Cheong, Nature **389**, 942 (1997).

A Selfconsistent-charge density-functional tight-binding scheme

M. Elstner[1,2], D. Porezag[1], G. Jungnickel[1], Th. Frauenheim[1], S. Suhai[2], and G. Seifert[3]

[1] *Technische Universität, Theoretische Physik III, D - 09107 Chemnitz, Germany*
[2] *German Cancer Research Center, Dept. Mol. Biophysics, INF 280, D-69120 Heidelberg*
[3] *Technische Universität, Institut für Theoretische Physik, D - 01062 Dresden, Germany*

ABSTRACT

We present an extension to the tight-binding (TB) approach to improve total energies, forces and transferability in the presence of considerable long-range Coulomb interactions. We derive an approximate energy expression in terms of charge density fluctuations δn at a reference (input) density n_0, which is a second order approximation to the total energy expression in density functional theory (DFT). With the choice of n_0 as a superposition of densities of neutral atomic fragments, we can define a repulsive potential as in standard TB theory, which is pairwise, short ranged and transferable. The zero order terms in the total energy expression are recoverd as the standard terms of our density-functional based tight-binding (DF-TB). For the second order terms, the charge density fluctuations δn are approximated by the total charge fluctuation Δq_α at atom α, which is qualitatively estimated by employing the Mullikan charge analysis. Within this approximations the total energy expression contains new parameters, which are related to ab-intio DFT calculations. Finally, by introducing localized basis functions and applying the variational principle we arrive at the Hamilton matrix elements, wich themselves depend on the charge fluctuations and, therefore, the general eigenvalue problem has to be solved self-consistently. To obtain forces for efficient geometry relaxation and molecular-dyamics, we calculated analytical derivatives of the total energy with respect to the atomic sites. In order to demonstrate the strenghts of our self-consistent-charge tight-binding (SCC-TB), we calculated reaction energies, geometries and vibrational frequencies for a large set of molecules and compare the results to semi-empirical methods, density-functional calculations and experiment.

I. INTRODUCTION

It has been shown that the TB-approach in general may be understood as a stationary approximation to density-functional theory (DFT) [1, 2, 3, 4, 5]. Central features of the common methodology, namely, non-selfconsistent treatment of the *Kohn-Sham* equations and the exploitation of pairwise repulsive interactions are strongly related to an appropriate "educated guess" for the inital charge density of the system.

In standard tight-binding theory, the calculated total energy differs from the true ground state energy in second order of the charge density fluctuations, which can be shown to be small for a properly chosen input density. For an input density, which is the superposition of the electronic density of neutral atomic fragments, a repulsive part in the total energy expression can be defined, which is pairwise, short ranged and depends only on the chosen input density [2]. In heteronuclear systems with considerable charge transfer, where the groundstate density may not be representable as a superposition of neutral atomic fragment densities, i.e. the initial charge density is not close to the ground state density, the second order corrections may become important. On the other hand, choosing an input density close to the true ground state density would inquire a priori knowledge of the particular system. Several proposals have been made to genearlize the TB-energies to account for charge transfer effects (see Frauenheim et al., this volume for references). But the need for a self-consistent extension of the TB-scheme is not necessarily indicated by a large charge transfer in the system. As Harrison has pointed out [7] for example, in really ionic systems with large charge transfer (e.g. NaCl) the non-selfconsistent methods lead to

131

an adequate description in many cases. The terms in a self-consistent TB method, which correct for the charge flow, tend to cancel in these cases. However, it is not valid in general and breaks down in systems with more subtle charge transfer effects, as for example many organic molecules, surfaces or defects in solids. Here, we focus on a systematic extension of the tight-binding formalism and of our DF-TB scheme [3, 6] in order to derive a generalized selfconsistent charge (SCC) methodology [13]. This differs from previous approaches since we base the modification of the TB total energy expression on a second-order expansion of the *Kohn-Sham* energy functional [8] with respect to density fluctuations.

II. SELFCONSISTENT-CHARGE TIGHT-BINDING METHODOLOGY

In the non-selfconsistent DFTB scheme the Hamilton matrix elements are calculated explicitely within an atomic orbital basis [6] and depend on the charge densities of neutral atomic fragments n_0. These scheme yields accurate results for a broad range of bonding situations, for which the superposition of overlapping atom-like densities serves as a good approximation for the many-atom structure. In order to recover the non-selfconsistent DFTB scheme as a zero order approximation for vanishing charge flow, we expand the *Kohn-Sham* energy expression up to second order at the input density n_0 (see Frauenheim et al., this volume), yielding:

$$E = \sum_i^{occ}\langle\Psi_i|\hat{H}_0|\Psi_i\rangle - \frac{1}{2}\iint'\frac{n_0'n_0}{|\vec{r}-\vec{r}'|} + E_{xc}[n_0] - \int V_{xc}[n_0]n_0 + E_{ii}$$

$$+\frac{1}{2}\iint'\left(\frac{1}{|\vec{r}-\vec{r}'|} + \left.\frac{\delta^2 E_{xc}}{\delta n\,\delta n'}\right|_{n_0}\right)\delta n\,\delta n'. \tag{1}$$

The traditional TB-approach is to preserve the first 5 terms in this final equation, with $\hat{H}_0$ as the Hamiltonian operator resulting from an input density n_0. The *Kohn-Sham* equations are then solved non-selfconsistently and the second-order correction is neglected. However, in a situation with considerable charge fluctuations, the second-order term in (1) may yield important corrections.

The contributions in (1) that depend on the input density n_0 only and the core-core repulsion are taken to be a sum of one- and two-body potentials [2]. So, denoting the latter by E_{rep}, we can define a repulsive Potential, which is strictly pairwise, repulsive and short-ranged also in the presence of considerbale charge flow. Since E_{rep} depends on n_0 only, this term is *not* influenced by the charge redistribution.

In order to include the effects of charge transfer in a simple and efficient TB concept, we first decompose $\delta n(\vec{r})$ into atom-centered contributions which decay fast with increasing distance from the corresponding center. The second-order term then reads:

$$E_{2nd} = \frac{1}{2}\sum_{\alpha,\beta}^{N}\iint'\Gamma[\vec{r},\vec{r}',n_0]\delta n_\alpha(\vec{r})\,\delta n_\beta(\vec{r}'), \tag{2}$$

where we have used the functional Γ to denote the *Hartree* and XC contributions. Second, the δn_α may be written as a product of radial and angular functions:

$$\delta n_\alpha(\vec{r}) = \sum_{l,m}K_{ml}F^\alpha_{ml}(|\vec{r}-\vec{R}_\alpha|)Y_{lm}\left(\frac{\vec{r}-\vec{R}_\alpha}{|\vec{r}-\vec{R}_\alpha|}\right)$$

$$\approx \Delta q_\alpha F^\alpha_{00}(|\vec{r}-\vec{R}_\alpha|)Y_{00}, \tag{3}$$

where F^α_{ml} denotes the normalized radial dependence of the density fluctuation on atom α for the corresponding angular-momentum. While the angular deformation of the charge

density, e.g. in covalently bonded systems, is already described very well within the non-SCF approach, the charge-transfer effects are not properly handled. Therefore, we truncate the multipole expansion (5) to the monopole term, accounting for the leading charge transfer contributions at conserved total charge, thus guaranteeing $\sum_\alpha \Delta q_\alpha = 0$. Substitution of (5) into (2) yields the simple final expression for the second-order energy term:

$$E_{2nd} = \frac{1}{2} \sum_{\alpha,\beta}^{N} \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} , \quad where \tag{4}$$

$$\gamma_{\alpha\beta} = \iint{}' \Gamma[\vec{r}, \vec{r}', n_0] \frac{F_{00}^\alpha(|\vec{r} - \vec{R}_\alpha|) F_{00}^\beta(|\vec{r}\,' - \vec{R}_\beta|)}{4\pi}. \tag{5}$$

is introduced as shorthand. In the limit of large interatomic distances, the XC contribution vanishes within LDA and E_{2nd} represents a pure Coulomb interaction between two point charges Δq_α and Δq_β. In the opposite case, where the charges are located at one and the same atom, a rigorous evaluation of $\gamma_{\alpha\alpha}$ would require the knowledge of the actual charge distribution. In order to avoid the numerical effort associated with such an approach, we assume that the self-interaction of the spherical charge fluctuation is essentially unaffected by the neighborhood and thus is the one of a free spin-unpolarized atom, i.e. it primarily depends on the atom type. This approximation is widely used in semi-empirical quantum chemistry methods relying on Parisers observation [9], that $\gamma_{\alpha\alpha}$ can be approximated by the difference of the atomic ionisation potential and the electron affinity, which is two times the chemical harndess η_α, or the Hubbard parameter U_α: $\gamma_{\alpha\alpha} \approx I_\alpha - A_\alpha \approx 2\eta_\alpha = U_\alpha$. The expression for $\gamma_{\alpha\beta}$ then only depends on the distance between the atoms α and β and on the so-called chemical hardness or Hubbard parameters U_α and U_β. These are constants and can be calculated for any atom type within LDA as the second derivative of the total energy in DFT with respect to the occupation number of the highest occupied atomic orbital.

Common functional forms for $\gamma_{\alpha\beta}$, which interpolate between these two cases, have been presented by Mataga-Nishimoto [12] and Dewar-Sabelli [10], Klopman [11]. In this work we use the latter which has the functional form

$$\gamma_{\alpha\beta} = \frac{1}{\sqrt{R_{\alpha\beta} + \frac{1}{4}(\frac{1}{U_\alpha} + \frac{1}{U_\beta})^2}}$$

Finally, the DFT total energy (1) is conveniently transformed into a transparent TB-form,

$$E = \sum_{i}^{occ} \langle \Psi_i | \hat{H}_0 | \Psi_i \rangle + \frac{1}{2} \sum_{\alpha,\beta}^{N} \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta + E_{rep}, \tag{6}$$

As discussed earlier, the contribution due to $\hat{H}_0$ depends only on n_0 and is therefore exactly the same as in the previous non-SCF studies [6]. However, since the atomic charges depend on the one-particle wave functions Ψ_i, a selfconsistent procedure is required to find the minimum of expression (6).

To solve the *Kohn-Sham* equations, the single-particle wave functions Ψ_i are expanded into a suitable set of localized atomic orbitals φ_ν for which the expansion coefficients are denoted by $c_{\nu i}$ [6]. By applying the variational principle to the energy functional (6), we obtain the *Kohn-Sham* equation which, within the pseudoatomic basis, transforms into a set of algebraic equations. For quantitatively estimating the charge fluctuations Δq_α, we employ the Mulliken charge analysis. Then, the *Kohn-Sham* equation reads:

$$\sum_{\nu}^{M} c_{\nu i}(H_{\mu\nu} - \varepsilon_i S_{\mu\nu}) = 0, \quad \forall \mu, i, \tag{7}$$

133

$$H_{\mu\nu} = \langle\varphi_\mu|\hat{H}_0|\varphi_\nu\rangle + \frac{1}{2}S_{\mu\nu}\sum_\xi^N(\gamma_{\alpha\xi} + \gamma_{\beta\xi})\Delta q_\xi \qquad (8)$$

$$= H_{\mu\nu}^0 + H_{\mu\nu}^1 \; ; \quad S_{\mu\nu} = \langle\varphi_\mu|\varphi_\nu\rangle \; ; \quad \forall\mu\epsilon\alpha, \; \nu\epsilon\beta.$$

Since the overlap matrix elements $S_{\mu\nu}$ generally extend over a few nearest neighbor distances, they introduce multiparticle interactions. The second-order correction due to charge fluctuations is now represented by the Mulliken charge dependent contribution $H_{\mu\nu}^1$ to the matrix elements $H_{\mu\nu}$.

As in previous studies [6], we have determined the short-range repulsive pair potential E_{rep} as a function of distance by taking the difference of the SCF-LDA cohesive energy and the corresponding TB band structure energy (which is now modified by the Coulomb correction) for a suitable reference structure. A simple analytic expression for the interatomic forces for use in efficient MD simulations is easily derived by taking the derivative of the final TB energy (6) with respect to the nuclear coordinates, see e.g. [14].

RESULTS

In the remainder of this paper, we shall concentrate on presenting results of first successful applications of the SCC-DFTB scheme to a wide range of systems, being of interest in chemistry, biology and physics. Thus, we demonstrate the improvements of the method as compared to conventional DF-TB, other semiempirical methods, *ab initio* calculations and experiment.

For our first benchmark, we have calculated the reaction energies of 36 processes between small closed shell molecules containing oxygen, nitrogen, carbon and hydrogen from Ref.[15], some of them shown in Tab. I.

Table I: Hydrogenation reactions (kcal/mol) for small organic molecules in comparison with DFT-LSD calculations and experiment [15].

reaction	SCC-TB	LSD	exp
$CH_3CH_3 + H_2 \rightarrow 2CH_4$	20	18	19
$CH_3NH_2 + H_2 \rightarrow CH_4 + NH_3$	23	24	26
$CH_3OH + H_2 \rightarrow CH_4 + H_2O$	32	28	30
$NH_2NH_2 + H_2 \rightarrow 2NH_3$	30	43	48
$HOOH + H_2 \rightarrow 2H_2O$	101	80	86
$CH_2CH_2 + 2H_2 \rightarrow 2CH_4$	71	67	57
$CH_2NH + 2H_2 \rightarrow CH_4 + NH_3$	66	67	64
$CH_2O + 2H_2 \rightarrow CH_4 + H_2O$	65	67	59
$NHNH + 2H_2 \rightarrow 2NH_3$	56	89	68
$C_2H_2 + 3H_2 \rightarrow 2CH_4$	124	131	105
$HCN + 3H_2 \rightarrow CH_4 + NH_3$	88	102	76
$CO + 3H_2 \rightarrow CH_4 + H_2O$	83	93	63
$N_2 + 3H_2 \rightarrow 2NH_3$	37	71	37

For all 36 reactions we have found a mean absolute deviation from experiment of 12.5 kcal/mol for the SCC-TB, compared to 11.1 kcal/mol for the DFT-LSD calculations. In further considering the optimized geometries of a 63 organic molecules test set from Ref [16], the mean absolute deviations in the bond lenghs and bond angles from experiment are $\Delta R = 0.010\mathring{A}$ and $\Delta\theta = 1.95°$ [14] respectively, evaluating 158 bondlenghts and 64 bond angles. This has to be compared to the deviations of other semiempirical methods like AM1 and others as listed in Tab. II, which has been taken from Ref. [21].

The values of the table are based on on a larger test set, evaluated for 228 bond lenghts and 92 bond angles. However, evaluating the errors of the AM1 Hamiltonian for

Table II: Performance of four different semiempirical methods for a test set of molecules, containing oxygen, nitrogen, carbon and hydrogen. Mean absolute errors, evaluated for 228 bond lenghts and 92 bond angles.

semiempirical Methods	ΔR ($\mathring{A}$)	$\Delta\theta$ (°)
MNDO	0.015	2.69
AM1	0.017	2.01
PM3	0.011	2.22
ECP-MNDO	0.012	1.88

the same test set as used for the SCC-TB, the AM1 errors appear even to be slightly worse as indicated in the Table. The improvement over the non-SCF treatment is impressively demonstrated for systems with a delicate counterbalance between ionic and covalent bonding contributions, as e.g. in formic (c.f. Table III).

Table III: Optimized geometries (bondlengths in $\mathring{A}$) of formic acid for the DF-TB, the SCC-TB , AM1 Methods and experiment).

	DF-TB	SCC-DFTB	AM1	exp.
C=O	1.282	1.215	1.230	1.202
C−O	1.292	1.349	1.356	1.343
O−H	0.984	0.978	0.972	0.972
C−H	1.100	1.119	1.103	1.097
OCO	116.7	123.1	117.7	124.9

The DF-TB clearly overestimates the equalization of the single and double bonds between carbon and oxygen. This is exclusively due to too much charge flow *(of nearly one electron)* from carbon to oxygen, clearly indicating the need for a selfconsistent charge redistribution.

Further, we have tested frequencies a series of 33 standard organic molecules yielding 6.4 % mean absolute deviation of vibrational frequencies from the experiment [14].

The very promising results for the organic molecules described above lead us to believe that the method can be applied to investigate the geometric and electronic structure of large biomolecules. In order to demonstrate this, we have simulated the retinal in the bacteriorhodopsin molecule, a polyene structure linked via a Schiff base to the protein α-Helix. The SCC-TB geometries are in good agreement with the experimentally reported crystalline structure of the retinal [17]. In particular, the planar relaxation of the retinal Schiff base is described correctly, a characteristic which both classical force field and semiempirical calculations have failed to model [18, 19]. In Tab. IV we show the calculated rotational barriers of some molecules in comparison with AM1 calculations and experiment. Although the barrier in ethylene is overestimated, all other barriers compare very well with experiment. This holds especially for the rotational barrier in formamide, which is a very important property for a realistic simulation of structural and energetic properties of peptides and proteins. This is so largely underestimated by semiempirical methods like AM1, that an additional correcting empirical force field must be introduced there for this purpose.

Further, the effects of hydrogen bonding within the SCC-TB are described properly, the correct geometry for the water dimer is predicted and the barrier for proton transfer in the $H_5O_2^+$ test-system is overestimated by only 12 % compared to DFT-B3LYP calculations. This opens the path for studies of the retinal in the protein enviroment and proton transfer mechanisms, including several hundred atoms of the retinal enviroment [20]. Since the time-limiting step in both, non-SCF and SCC-TB calculations is the solution of the general

Table IV: Rotational barriers (kcal/mol) for small organic molecules in comparison with semi-empirical AM1 [16] calculations and experiment [16].

	SCC-TB	AM1	exp
C_2H_4	100	65.0	65.9
C_2H_6	2.4	1.25	2.9
CH_3NH_2	2.25	2.29	2.0
CH_3OH	1.3	1.04	1.1
formamide	18.3	10.11	≈ 20

eigenvalue problem, one step in the MD-simulations for structural relaxations by means of the SCC-TB requires this problem to be solved 2 to 3 times on average, independent on implementation.

SUMMARY

Summarizing, we have presented a selfconsistent-charge extension of the non-orthogonal *(two-center)* DF-TB scheme based on a second-order expansion of the total energy as calculated within DFT. The method can be seen as a general SCC-extension of TB theory offering the great advantage to incorporate any atom type in a straightforward manner. This hopefully will stimulate not only MD-applications to large-scale semiconductor structures and biological systems, but also for other challanging types of materials.

REFERENCES

1. J. Harris, Phys. Rev. **31** 1770 (1985).
2. W. Foulkes, R. Haydock, Phys. Rev. B **39** 12520 (1989).
3. G. Seifert, H. Eschrig, W. Bieger, Z. Phys. Chemie *(Leipzig)* **267** 529 (1986).
4. A. P. Sutton, et al., J. Phys. C: *Solid State Physics* **21** 35 (1988).
5. O. F. Sankey, D. J. Niklewski, Phys. Rev. B **40** 3979 (1989).
6. D. Porezag, Th. Frauenheim, Th. Köhler, G. Seifert, R. Kaschner, Phys. Rev. B **51** 12947 (1995).
7. W. A. Harrison, Phys. Rev. B **31** 2121 (1985).
8. W. Kohn, L. J. Sham, Phys. Rev. **140A** 1133 (1965).
9. R. Pariser, J. Chem. Phys. **24** 125 (1956)
10. M. J. S. Dewar, N. L. Hojvat (Sabelli), J. Chem. Phys., **34** 1232 (1961), M. J. S. Dewar, N. L. Sabelli, J. Phys. Chem., **66** 2310 (1962),
11. G. Klopman, J. Am. Chem. Soc. **28** 4550 (1964).
12. N. Mataga, K. Nishimoto, Z. phys. Chemie (Frankfurt), **13** 140 (1957).
13. M. Elstner, et al., submittet to PRL.
14. M. Elstner, et al., to be published in Phys. Rev. B.
15. J. Andzelm, E. Wimmer, J. Chem. Phys. **96** 1280 (1992).
16. J. S. Dewar et al., J. Am. Chem. Soc. **107** 3902 (1985)
17. B. Santarsiero et al., J. am. Chem. Soc. **112** 9416 (1990).
18. E. Tajkhorsid, B. Paiz, S. Suhai, J. Phys. Chem. **1997** in press.
19. B. Tavan, K.Schulten, D. Osterhalt, Biophys. J. **47** 415 (1985).
20. E. Tajkhorsid, M. Elstner, T. Frauenheim , S. Suhai, to be published.
21. M. Kolb, W. Thiel, J. Comp. Chem., **14** (1993) 775

TIGHT-BINDING LINEAR MUFFIN-TIN ORBITAL IMPLEMENTATION OF THE DIFFERENCE EQUATION GREEN'S FUNCTION APPROACH FOR 2D-PERIODIC SYSTEMS

MARK VAN SCHILFGAARDE[a] AND WALTER R. L. LAMBRECHT[b]
[a] SRI, INTERNATIONAL, 333 RAVENSWOOD AVENUE, MENLO PARK, CA 94025
[b] DEPARTMENT OF PHYSICS, CASE WESTERN RESERVE UNIVERSITY, CLEVELAND, OH 44106-7079

ABSTRACT

The central ideas and the advantages of the difference equation Green's function approach for layered systems with 2D-periodicity are presented. The tight-binding linear muffin-tin orbital and other practical aspects of the implementation are discussed. Preliminary test results are presented.

INTRODUCTION

Layered materials with 2D-periodicity form an important class of systems for electronic structure studies, warranting a special and optimized treatment. They include surfaces, interfaces, single and multiple quantum well heterostructures. Such systems are important for metals (e.g. magnetic layered structures), semiconductors, and for combinations of the two (e.g. Schottky barriers, metal based resonant tunneling structures). In this paper, we present the key ideas and preliminary results of a new method, which makes optimal use of the layered arrangement. The method is based on the difference equation approach (DEA) of Chen et al. [1] and implemented in the framework of the tight-binding linear muffin-tin orbital (TB-LMTO) method [2] in its atomic sphere approximation (ASA). This makes the method compatible with the local density functional approach and as such a first-principles method. It is closely related to the LMTO surface Green's function method (GFM), which was proposed in Ref. [3] and subsequently developed and used successfully by Skriver et al. [4, 5, 6] and Kudrnovský et al. [7].

While the current restriction to ASA limits its applicability to solve structural problems, it is hoped that future developments of full charge density LMTO [8] within the same general theoretical framework, will allow to overcome this difficulty. The main strength of the presently proposed method is that it makes explicit use of the fact that when several layers (to be defined properly below) see effectively the same potential, the DEA allows us to "jump" over these layers. This implies that the method scales effectively with the number of interfaces rather then with the number of atoms. At the heart of the method is the observation that charge density and potential perturbations are usually confined to the immediate neighborhood of each interface, while individual wave functions or the Green's function (GF) at a particular energy are much longer range. As such, the Green's functions propagating through these bulk-like regions can still establish important interaction effects between interfaces. We anticipate that the method will thus be most useful for problems which involve some longer range interaction effect between rather well separated interfaces. Examples of this are oscillatory magnetic couplings in metallic artificially layered structures and transport in resonant tunneling structures in semiconductors. In both these examples, the physics involves primarily electrons near the Fermi energy. We

emphasize that the central quantities in both these problems involve off-diagonal elements of the Green's function. Because the DEA obtains these exactly for any stack of identical layers in terms of the complex band structure [9], no numerical accumulation of errors takes place when increasing the distance which connects the off-diagonal elements.

At the same time, the potential variations near each interface must be obtained self-consistently. This involves primarily the diagonal elements of the Green's function because they determine the local charge density. We show explicitly that this can be done in order-N matrix operations with N the number of layers. In order to dampen the charge sloshing effects, which considerably slow convergence to self-consistency, especially in large systems, one may dramatically accelerate convergence by estimating the self-consistent density from the input and output density in the linear response approximation [10, 5]. Entering into the linear response is the unscreened susceptibility matrix, which tells us how much charge is induced at one site when the potential is changed at another site. It is effectively obtained in terms of the off-diagonal Green's functions as

$$\chi_{RL,R'L'} = \frac{1}{\pi} \mathrm{Im} \oint dz \sum_{k_\parallel} G^0_{RL,R'L'}(z, k_\parallel) G^0_{R'L',RL}(z, k_\parallel), \tag{1}$$

in which the GF and the energy contour will be specified below. This emphasizes again the importance of off-diagonal GF matrix elements, for which the DEA is particularly suited.

TB-LMTO ASPECTS

The key feature of any tight-binding method is the range of the interactions. In the present context of layered systems, it defines the size of the so-called principal layer (PL). This is a layer of such size that interactions couple only nearest neighbor PLs. A 2D Bloch-transformation exploiting the 2D periodicity is tacitly understood. By this partitioning in PLs, the electronic structure problem is then reduced to matrix equations which are block tridiagonal. As will be shown below, the DEA as other surface GFMs exploit this sparseness of the matrix equations optimally.

Within the ASA-LMTO framework, the multiple-scattering or "tail-cancellation" equations provide the most suitable formulation of the electronic structure problem for the present purpose. The associated Green's function equation is:

$$[P_{RL}(z)\delta_{RL,R'L'} - S_{RL,R'L'}(k_\parallel)]g_{R'L',R''L''}(z, k_\parallel) = \delta_{RL,R''L''}, \tag{2}$$

with $P_{RL}(E)$ the potential functions (essentially cotangent of the phase shift), and $S_{RL,R'L'}$ the structure constants. Using the TB screening representation of the LMTO method [2], the structure constants are typically limited to next nearest neighbors while the Hamiltonian matrix involves products of such structure constants and thus reaches at least to 4th nearest neighbors. Additionally, when the potential is perturbed on site R, only the local potential function is changed, whereas off-diagonal Hamiltonian matrix elements to the neighboring atoms are also changed. For the present purpose it is sufficient to note that the full GF in real space, i.e. $G(\mathbf{r}, \mathbf{r}', z, k_\parallel)$ can be obtained from it, as well as all other necessary quantities for a self-consistent calculation. For example, the local charge density in angular momentum channel L at site R is given by

$$n_{RL} = \frac{1}{\pi} \mathrm{Im} \oint dz \sum_{k_\parallel} \{[\dot{P}_{RL}(z)]^{1/2} g_{RL,RL}(z, k_\parallel)[\dot{P}_{RL}(z)]^{1/2} - d\ln[\dot{P}_{RL}(z)]^{1/2}/dz\}, \tag{3}$$

in which the contour is a semicircle in the lower half plane starting below the bands of interest and cutting into the real axis at the Fermi level and the dot means energy derivative.

PRINCIPLE LAYER GREEN'S FUNCTIONS

As already mentioned, the PL technique, which is central to all TB surface GFMs, exploits the property that the $(P - S)$ matrix is tridiagonal in the PL-representation. One can build up the GF from left to right by adding layer by layer [7]. Each step corresponds to applying the Dyson coupling equation:

$$[P_n - S_{nn} - S_{nn-1}g_{n-1n-1}^{(n-1)}S_{n-1n}]g_{nn}^{(n)} = 1_{nn}, \tag{4}$$

in which $g_{nn}^{(n)}$ is the GF of the n-th layer when the system is truncated at that layer, S_{n-1n} is the structure constant matrix coupling layer $n - 1$ to layer n, the energy and $k_{\parallel}$ arguments have been suppressed, and the elements are understood to be matrix blocks labeled by sites and orbitals in each layer. For an interface system, one starts from the surface GF block of an ideal semi-infinite crystal (which can be obtained in a variety of ways, e.g. see below), and in the last step, one couples simultaneously the last layer of the interface region (say, layer N) to the stack of layers on the left and the surface block of the semi-infinite ideal crystal on the right by adding the term $-S_{NN+1}g_{N+1N+1}^{N+1}S_{N+1N}$ to the expression in square brackets in Eq. 4. At this point one has the true GF of the interface system on the last layer N. One can then retrace one's steps to update the GF of the previous layers in the stack. It is important to realize that at each step, the true GF g is known on the previous layer coming from the right (say $n+1$), as well as the GF of the uncoupled system terminated on the current layer n, $g_{nn}^{(n)}$. One obtains from the Dyson equation

$$
\begin{aligned}
g_{nn+1} &= g_{nn}^{(n)}S_{nn+1}g_{n+1n+1}, \\
g_{n+1n} &= g_{n+1,n+1}S_{n+1n}g_{nn}^{(n)}, \\
g_{nn} &= g_{nn}^{(n)} + g_{nn}^{(n)}S_{nn+1}g_{n+1,n}.
\end{aligned}
\tag{5}
$$

The computational effort required to obtain all diagonal GF blocks in the interface layer scales linearly in the number of layers, N. (Inversions and multiplications take comparable computing time.) Additional off-diagonal Green's functions can be similarly computed: a band of $g_{n-m,n+m}$ requires order mN operations for N layers. For a straightforward self-consistent calculation, only the diagonal matrix elements are required. The off-diagonal matrix elements decay exponentially as the distance between the atoms increases; thus one needs only a few off-diagonal elements to generate, for example, χ.

DIFFERENCE EQUATION GREEN'S FUNCTION APPROACH

The DEA [1] takes advantage of the fact that the various matrix blocks occuring in the secular equations are the same whenever they apply to bulk-like layers. Let us reconsider the secular equation for a stack of equivalent layers, i.e. layers with the same potential and structure. Introducing the notation $S_{10} \equiv S_{n+1,n}$, $(S-P)_{00} \equiv S_{nn} - P_n$, (which emphasizes that these are independent of the actual layer n), the equation becomes:

$$S_{10}^{\dagger}C_{n-1} + (P - S)_{00}C_n + S_{10}C_{n+1} = 0 \tag{6}$$

Because this is a second order difference equation in the layer index, its general solution can be written as $C_n = \sum_j a_j r_j^n V_j$ in terms of the eigenvalues r_j and eigenvectors V_j of the quadratic characteristic equation:

$$[S_{10}^\dagger + (S - P)_{00} r + S_{10} r^2] V = 0, \tag{7}$$

The eigenvalues of this equation occur in pairs $r, 1/r$; this is a consequence of Eq. 7 being a quadratic eigenvalue problem. Their physical meaning is that $r = \exp(i k_z a_z)$ with a_z the distance between the layers and k_z a complex wave vector, provide the so-called complex band-structure. They are equivalent to the complex wave numbers in the eigenvalue method of Chang and Schulman [9]. The GF g_{mn}^B satisfies the same equation as C_n when $m \neq n$. When $m = n$, the right hand side must be replaced by -1_{nn}. If $m \geq n$, we can write

$$g_{mn}^B = V_< r^{m-n} V_<^{-1} g_{nn}, \tag{8}$$

in which r is now a diagonal matrix containing only the eigenvalues for which $|r| < 1$ and $V_<$ is a square matrix obtained by juxtaposing the corresponding eigenvector columns because g_{mn} must decrease with increasing distance. Similarly, if $m \leq n$, only eigenvalues $|r| > 1$ appear, which correspond to $1/r$ in the previous set. (By adding an imaginary part to the energy, one insures that the eigenvalues strictly fall in either one of these two categories.) The equation for the bulk GF then becomes:

$$[S_{10}^\dagger V_> r V_>^{-1} + (S - P)_{00} + S_{10} V_< r V_<^{-1}] g_{nn}^B = -1. \tag{9}$$

So, the bulk GF is obtained by one inversion once the characteristic eigenvalue problem has been solved. The left and right ideal semi-infinite GF, i.e. the GF for the truncated systems, can be obtained either from

$$[S_{10}^\dagger V_> r V_>^{-1} + (S - P)_{00}] g_{nn}^{LS} = -1 \tag{10}$$

$$[(S - P)_{00} + S_{10} V_< r V_<^{-1}] g_{nn}^{RS} = -1 \tag{11}$$

in which only one "coupling term" appears, or, using the Löwdin downfolding approach [11, 12], in which the SGF is given by

$$g_{mn}^S = g_{mn}^B - g_{m0}^B [g_{00}^B]^{-1} g_{0n}^B. \tag{12}$$

Since $[g_{00}^B]^{-1}$ as well as the off-diagonal g_{m0}^B are already known from Eqs.(9,8), we have for $m \geq n$,

$$g_{mn}^S = \{[V_< r^{m-n} V_<^{-1}] - [V_> r^m V_>^{-1}][V_< r^n V_<^{-1}]\} g_{00}^B. \tag{13}$$

The important thing to realize is that Eq. 8 can be used to obtain the required off-diagonal elements of the GF linking the blocks on either end of a stack of equal layers without the need to determine the ones in between. This is because the information on the evanescent and propagating states is already contained in the complex band structure eigenvalues and eigenvectors. As Eq. 13 illustrates, any surface GF and, in fact, any finite piece GF can be written analytically in terms of r, $V_<$, $V_>$ and the bulk diagonal g_{00}^B. It means that the computational effort is independent of how many equivalent layers occur. In other words, the method becomes order-N in the number of interfaces, rather than the number of layers.

TEST RESULTS

Because of the central role played by the complex band structure in the DEA, we show an example the complex band structure of Si along the [110] direction in Fig. 1. Good agreement is obtained on the topology of the bands with previous pseudopotential calculations of Chang and Schulman.[9]
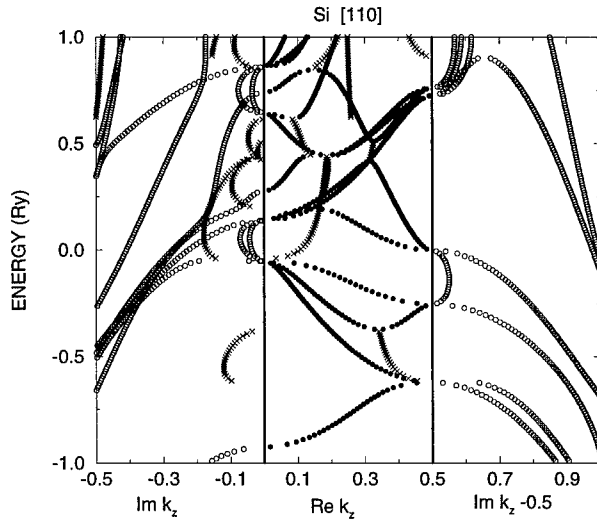
Figure 1: Complex band structure of Si [110]. Real bands are indicated by filled circles, imaginary bands with real part zero by open circles with negative Im k_z, imaginary bands with real part at the zone boundary by open circles with positive Im k_z and complex bands by ×.

An important test is whether the method provides the same result as a standard band structure calculation in case of a bulk periodic system. Fig. 2 shows the density of states of bcc Fe obtained with the present GFM and from the band structure using the tetrahedron method. The GFM uses a 16 × 16 point 2D k-mesh. In fact, the calculation was made self-consistently entirely using the GFM and converges to the same result. One important difference with bulk band structure calculations is that in GF calculations the Fermi level is fixed rather than the number of electrons in the cell. Therefore, one must shift the potential by a constant until the correct number of electrons required by charge neutrality is obtained. In Fe, this requires high accuracy on the DOS, because of the sharp peak at the Fermi level, thus providing a demanding test.

CONCLUSIONS

The difference equation Green's function approach is a promising technique for 2D periodic systems. It makes use of the complex band structure. Preliminary tests of its TB-LMTO implementation are satisfactory. Applications are under development.
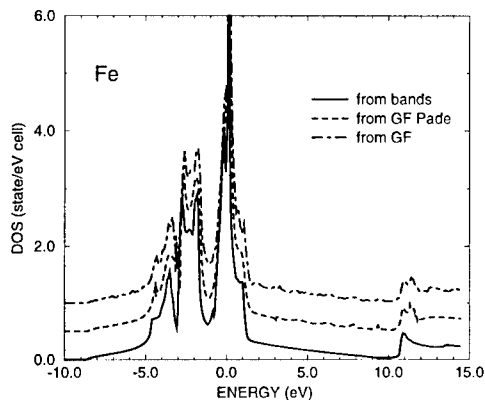
Figure 2: Density of states in bcc Fe. Solid line: standard LMTO band structure calculation; dot-dashed line: GFM along a line just below the real axis; dashed line: GFM using Padé approximation to the GF along an elliptical contour in lower half plane (using only 40 complex z-points) and analytic continuation. The various curves are offset for clarity. The Fermi energy is chosen as zero of the energy scale.

REFERENCES

1. A. B. Chen, Y. M. Lai-Hsui, and W. Chen, *Phys. Rev. B* **39**, 923 (1989).

2. O. K. Andersen and O. Jepsen, *Phys. Rev. Lett.* **53**, 2571 (1984).

3. W. R. L. Lambrecht and O. K. Andersen, *Surf. Science* **178**, 256 (1986); and unpublished notes.

4. H. L. Skriver and N. M. Rosengaard, *Phys. Rev. B* **43**, 9538-49 (1991)

5. H. L. Skriver and N. M. Rosengaard, *Phys. Rev. B* **46**, 7157-68 (1992).

6. N. M. Rosengaard and H. L. Skriver, *Phys. Rev. B* **47**, 12865 (1993);

7. J. Kudrnovský, V. Drchal, I. Turek, and P. Weinberger, *Phys. Rev. B* **50**, 16105-8 (1994).

8. O. K. Andersen, this MRS Proc. and O. K. Andersen, O. Jepsen and G. Krier, in *Lectures in Methods of Electronic Structure Calculation*, (World Sci. Publ., Trieste 1994).

9. Yia-Chung Chang and J. N. Schulman, *Phys. Rev. B* **25**, 3975 (1982).

10. W. R. L. Lambrecht, B. Segall, and O. K. Andersen, *Phys. Rev. B* **41**, 2813-31 (1990).

11. P. O. Löwdin, J. Chem Phys. **19**, 1396 (1951);

12. A. R. Williams, P. J. Feibelman, and N. D. Lang, *Phys. Rev. B* **26**, 5433 (1982).

INTERACTIONS OF POINT AND EXTENDED DEFECTS IN STRUCTURAL INTERMETALLICS: REAL-SPACE LMTO-RECURSION CALCULATIONS

O.YU.KONTSEVOI *, O.N.MRYASOV *, YU.N.GORNOSTYREV **, A.J.FREEMAN *

*Department of Physics and Astronomy, Northwestern University, Evanston, IL, 60208-3112.
**Institute of Metal Physics, Ekaterinburg, Russia

ABSTRACT

A real-space TB-LMTO-recursion method for electronic structure calculations is applied to the study of interacting extended and point defects in NiAl. Results of calculations for the pure intermetallic and with ternary additions (within a supercell model) show good agreement with band structure results. Further, electronic structure and total energy calculations of point (single impurity, M=Ti, V, Cr, Mn, Fe and Co) and planar defects such as anti-phase boundaries (APB) were carried out and the interaction between them was determined. We found that for the $\frac{1}{2}\langle 111\rangle\{110\}$ APB in NiAl, ternary additions occupy exclusively the 3d-metal sublattice and decrease the APB energy (except for Co). Finally, we employ TB-LMTO-REC to study the electronic structure of the most complex extended defect, a dislocation. We demonstrate for the $\langle 100\rangle\{010\}$ edge dislocation in NiAl that: (i) quasi-localized states may exist as a result of specific lattice distortions in the dislocation core with a type of "broken" bonds; (ii) the electronic structure changes appreciably in the process of dislocation motion; (iii) van-Hove singularities present in the ideal crystal may be shifted to E_F as a result of the dipolar character of the deformations in the dislocation core.

INTRODUCTION

Transition metal aluminides, such as NiAl, FeAl and γ-TiAl, are of great interest as structural materials. The mechanical properties of intermetallic compounds strongly depend on the presence of ternary additions, which may be able to change such characteristics as ductility, fracture mode, oxidation behavior, etc. Experiments show that small amounts of ternary additions in NiAl lead to increased ductility (solid solute hardening) [1]. The influence of impurities on the electronic structure and energetics of extended defects in crystals is of special interest in this context. It was shown previously [2] that the energy of anti-phase boundaries (APB) is the main factor which determines the preference slip systems in intermetallics. Alloying can change the APB energy and so can influence their mechanical behavior. Another mechanism, responsible for solid solute hardening, may be impurity-dislocation interactions. Theoretical investigations of such interactions by standard band structure methods are complicated due to the absence of translation symmetry. In this work we demonstrate that an alternative, real-space technique is the effective tool for studying the interactions of point and extended defects in structural intermetallics.

METHOD

For the electronic structure calculations, we employ the real-space tight-binding linear muffin-tin orbital method [3] combined with recursion (TB-LMTO-REC) [4], [5], tailored to treat an impurity problem as described in [6]. Perturbations of the electronic structure, introduced by point (impurity) or extended defects (APB or dislocations) can be treated rigorously within the region of several shells of nearest neighbors around the defect. This region is subsequently embedded in a larger cluster of several thousand atoms, which repre-

sents the bulk material, and then boundary conditions relevant to the problem are simulated. For the embedded region, the effects of the electron density redistribution is described self-consistently in the framework of the local density approximation and local charge neutrality within the atomic spheres of varying size is achieved. The perturbation of the remaining part of the cluster is taken into account approximately due to changes of structure constants under lattice deformations, but with LMTO potential parameters taken to be fixed as in the bulk. In the present work, all calculations were carried out in a basis of s, p and d states, for which, respectively, 10, 20 and 30 pairs of recursion coefficients were calculated. The procedure of Beer and Pettifor [7] was exploited for the correct termination of continued fractions. The second order Hamiltonian and exchange-correlation potential of von Barth and Hedin were used.

RESULTS

The precision of our implementation of the real space TB-LMTO-REC method and the self-consistency procedure employed was checked by calculating the equilibrium volume of pure NiAl. We found that the equilibrium lattice constant (a=2.824 Å) is in good agreement both with experiment (2.887 Å) and with results of band structure methods (FP-LMTO a=2.839 Å[2], FLAPW a=2.810 Å[8]). A 2% underestimation of the equilibrium lattice constant is typical for methods using the local-density approximation. We also found that the variations of recursion scheme parameters (cluster size from 1000 to 10000 atoms, length of continued fractions from 20 to 50, presence or absence of periodic boundary conditions) lead to less than 0.5 mRy changes in the self-consistent total energy, and do not change the position of theoretical equilibrium volumes.

To better understand the mechanisms of solution strengthening in ordered alloys, the information regarding the crystallographic site occupancy of the ternary element is required. To verify the applicability of the method for preferred site energy evaluation, we calculated its values for transition metals (TM) from Ti to Ni modeled with a small 16-atom supercell, as was done previously by Medvedeva, et.al. [9] by the FLMTO method, which corresponds to a 12.5% impurity concentration. These supercells form a cubic cluster of 2000 atoms with periodic boundary conditions. The preferred site energy is determined as the difference between the total energies of supercells with additions in both sublattices and of pure Ni and Al metals in their structural forms having the lowest total energy. The results are presented in Fig. 1 in comparison with results of [9]. One can see the very good agreement between real space and band supercell results.

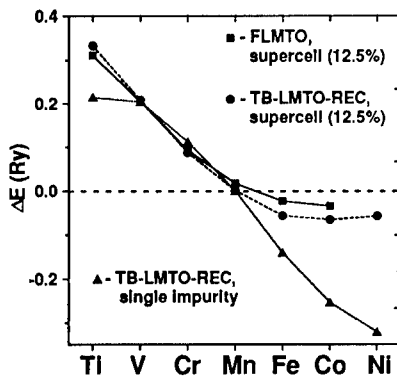Figure 1: **Preferred site energies for TM impurities in NiAl calculated by FLMTO and TB-LMTO-REC methods in supercell and single impurity approaches.**

According to our calculations, the site preference for Ti, V and Cr substitutions must be the Al sublattice, in agreement with experimental data [10]. For the series of transition metals, the preferred site energy decreases from Ti to Mn, while Fe and Co impurities have

a trend to replace Ni. For the Fe impurity this is in contradiction with experimental studies [10], which show for nickel-rich (with 0.3 at.% Fe) and for aluminum-rich (with 2.2 at.% Fe) NiAl alloys that the Fe substitutes Al in both alloys. The reason for discrepancies between theoretical and experimental results may be due to different impurity concentrations. The calculations were performed for the case of periodically ordered impurities with relatively high concentration, but it is known [1] that the solid solution hardening effect takes place at low impurity concentrations. To find out concentration differences in preferable site energies, we carried out the calculations for the case of infinitely dilute alloys, which were simulated by a single impurity placed in the center of 2000 atom clusters of NiAl. Potentials of the impurity atom and 6 shells of its nearest neighbors were included in the self-consistent procedure, giving us a cluster of 65 atoms embedded in a non-perturbed host. As shown in [6], in metallic systems six shells of nearest neighbors essentially screen the perturbation induced by the impurity atom to the host, and we have used changes of the total energies for the embedded cluster in estimations of the preferred site energies. These results are also presented in Fig. 1. We see that, for the case of a single impurity, a similar trend in preference energy and site is observed: it decreases from Ti to Mn, then the preference site is changed from Al to Ni. But for the single Fe impurity we have obtained an even higher preference energy for the Ni site than for the case of finite impurity concentration. There may be several reasons for such a discrepancy with experimental results: (i) lattice relaxations around the impurity were neglected in this work; (ii) experimental results were obtained for non-stoichiometric compositions, whereas it was shown in [12] that small deviation from stoichiometry may significantly change alloy formation energies, such as enthalpy of formation.
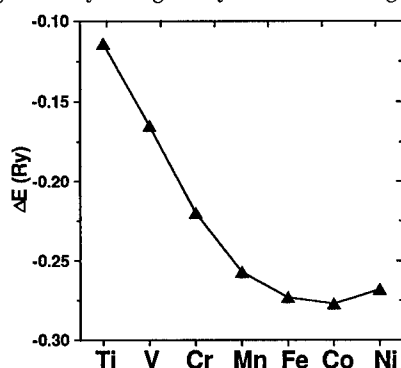
Figure 2: **Preferred site energies for TM impurities in $\frac{1}{2}\langle 111\rangle\{110\}$ APB in NiAl.**

According to both experimental [13] and theoretical [2] results, the major deformation mode in NiAl is $\langle 100\rangle\{110\}$ slip. Now, promotion of $\langle 111\rangle\{110\}$ slip through alloying may be promising for increasing the ductility of NiAl. We have investigated the influence of TM impurities on the $\frac{1}{2}\langle 111\rangle\{110\}$ APB energy in NiAl. In contrast to previous band structure calculations of APB's, we do not have to introduce a sequence of repeated supercells thereby avoiding possible interactions between periodic APB's in different supercells. In this work, atomic relaxations at the APB interface were neglected.

We have checked the dependence of the APB energy on the geometry of the cluster and have found that for clusters of 1300, 4200 and 10000 atoms the APB energy was 890, 920 and 920 mJ/m², respectively. It is seen that even a relatively small cluster of 1300 atoms gives an almost converged APB energy. Our results for the APB energy are also in good agreement with results obtained previously by band structure methods for an unrelaxed APB: 880 mJ/m² ([11], LMTO-ASA), 1000 mJ/m² ([2], FP-LMTO), 1000 mJ/m² ([12]). As shown in [8],[12], relaxation effects may result in a lowering of the APB energy by up to 20% of the unrelaxed value.

To study the interaction of transition metal impurities with the APB, a single impurity was placed in the APB plane, substituting either Al or Ni. Potentials of impurity atom and

three shells of its nearest neighbors were calculated self-consistently. The results of preferred site estimates are presented in Fig. 2. It was found, in contrast to the case of pure NiAl, that the preferred site for all transition metal impurities is the Ni site. The main reason for that is the change in impurity-nearest neighbor interactions, mainly due to changes of first neighbor position occupants.
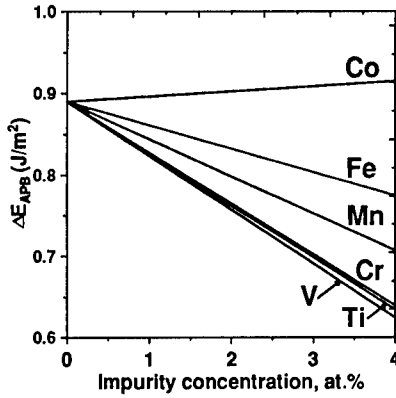
Figure 3: **Dependencies of APB energy on impurities concentration.**

Knowing the changes of total energies of the APB due to introduction of a single impurity when it is placed in the preferred position, we can evaluate the concentration dependence of the APB energy. Such an analysis is possible only for low impurity concentrations, and was made assuming the following approximations: (i) impurities do not interact with each other; (ii) the concentration of impurities on the APB is equal to those in the crystal (no impurity segregation on APB); (iii) only impurities in the APB plane can influence its energy. Taking into account that in our calculations perturbations due to the impurity were considered as localized within 3 neighbor shells, this "non-interacting limit" can be roughly estimated as about 4 at%. The calculated dependencies of the APB energy as a function of impurity concentration were calculated and presented in Fig. 3. Transition metal impurities cause the notable decrease of APB energy (up to 30% in the concentration region considered). The largest effect is for early impurities (Ti-Cr) and then decreases. By lowering the $\frac{1}{2}\langle 111\rangle\{110\}$ APB energy, early TM impurities promote $\langle 111\rangle\{110\}$ slip. However, to answer the question whether this effect is large enough to change the preferred slip system, a further detailed analysis of impurity-induced changes of deformation modes, dislocation structures and Peierls stress [2] is necessary.

Another factor, which may be involved in solid solution hardening, is impurity-dislocation interactions. In [14], in the framework of a semi-empirical model, the interactions of impurities with different types of dislocations were analyzed and it was shown that extra solid solution hardening may arise in cases when localized electronic states appear on dislocations. This is most probable for dislocations with $\langle 100\rangle$ Burgers vector.

We have investigated the electronic structure of an $\langle 100\rangle\{010\}$ edge dislocation. The coordinates of atoms in the dislocation core were determined within the Peierls-Nabarro model [14] with parameters obtained from ab-initio calculations of the generalized stacking fault energetics [2]. The effects of the electron density redistribution in the region with up to 100 atoms in the dislocation core (shown in Fig. 4) are treated self-consistently. The size of the cluster was taken to be about 10000 atoms in order to minimize the effects of the boundaries. Two non-equivalent dislocation cores are possible in NiAl depending on the position of the dislocation axis: with the top layer (open circles in Fig. 4) filled with (i) Ni atoms (which we denote as position A), and (ii) with Al atoms (which we denote as position B). Numbered atoms represents regions of compression (2-4, 9-11) and expansion (5-7, 12-14) within the dislocation core.

The local densities of states (LDOS) for Ni atoms in the center of the dislocation core are

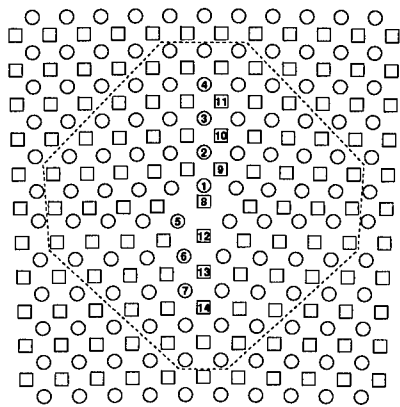presented in Fig. 5 for positions A and B, respectively. The main features of the electronic structure are:

Figure 4: **Positions of atoms in the** $\langle 100 \rangle \{010\}$ **edge dislocation core with the region marked by dashed lines treated self-consistently. Open circles denote atoms in top layer and open squares - in bottom layer.**

(i) well-defined peaks of quasi-localized states are seen for atoms in the center of the dislocation core. Hence, this effect may be associated with "broken" bonds or a change of coordination number in comparison with the bulk; it is most pronounced in the top layer, where distortions in the center of the dislocation core may be associated with "missing" atoms [14]; (ii) some features of the electronic structure, especially those corresponding to quasi-localized states for a dislocation in positions A and B are seen to be appreciably different. Hence, in the process of dislocation motion, the dislocation core introduces different perturbations into the electronic subsystem. (iii) The electronic structure of bulk NiAl is characterized by a van-Hove singularity below the Fermi level. Thus, there is the possibility that dipolar dilatation (compression and expansion regions in the dislocation core) may result in a shift of the vHS onto E_F for some atoms in the dislocation core. Our ab-initio calculations do support such a scenario for some atoms in the compression region of the dislocation core (atoms 2–4). This result may be treated as clear theoretical evidence for specific mechanisms of electron scattering by dislocations, which may possibly be revealed in the electronic transport properties of real metallic crystals.
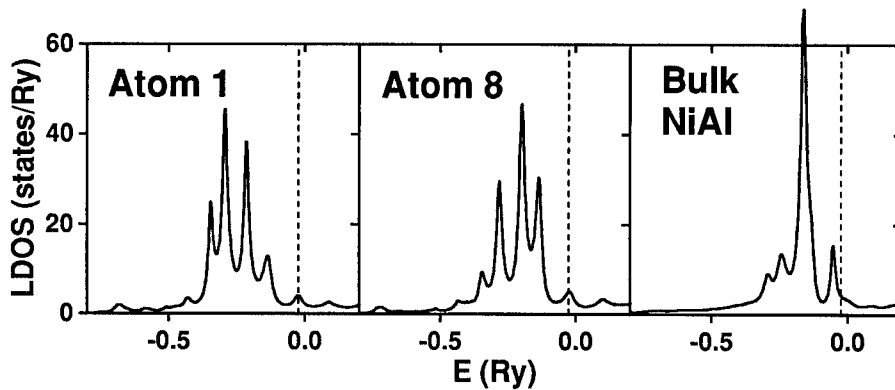

Figure 5: **Local densities of states for Ni atoms from the center of dislocations (atom 1 in the position A and atom 8 in the position B) in comparison with those in pure NiAl.**

CONCLUSIONS

We have demonstrated that, due to its real-space nature, the TB-LMTO-REC method allows one to evaluate the electronic structure in complex situations, typical for real materials, such as dislocations, planar defects and their interactions with point defects. For supercell calculations, the results from this method are in good agreement with conventional band structure results. We find that interactions of TM impurities with the $\frac{1}{2}\langle 111\rangle\{110\}$ APB in NiAl decrease its energy, and so may favour $\langle 111\rangle\{110\}$ slip. Further, the method allows one to address such extremely difficult problems as electronic states on dislocations. For the first time within an ab-initio method, we provide evidence for the possibility of electronic state localization in metallic systems due to the lattice distortion in the edge dislocation core.

ACKNOWLEDGMENTS

REFERENCES

1. R.D. Noebe, R.R. Bowman and M.V. Natal, International Materials Reviews **38**, p. 34 (1993).
2. N.I. Medvedeva, O.N. Mryasov, Yu.N. Gornostyrev, D.L. Novikov and A.J. Freeman, Phys.Rev. B **54**, p. 13506 (1996).
3. O.K. Andersen, O. Jepsen and D. Glötzel in Highlights of Condensed Matter Theory, edited by F. Bassani, F. Fumi and M. Tosi (Amsterdam, North-Holland, 1985), p. 59.
4. H.J. Nowak, O.K. Andersen, T. Fujiwara, O. Jepsen and P. Vargas, Phys.Rev.B **44**, p. 3577 (1991).
5. O.Yu. Kontsevoi, O.N. Mryasov and V.A. Gubanov, Fiz.Tverd.Tela, **34**, p. 2624 (1992) [Sov.Phys.Solid State **34**, p. 1506 (1992).]
6. O.Yu. Kontsevoi and V.A. Gubanov, Phys.Rev.B, **51**, P. 15125 (1995).
7. N. Beer and D.G. Pettifor in Electronic Structure of Complex Systems, edited by P. Phariseau and W.M. Temmerman (New York, Plenum, 1984), p. 769.
8. R. Wu, L. Zhong, L. Cheng and A.J. Freeman, Phys.Rev.B, **54**, p. 7084 (1996).
9. N.I. Medvedeva, Yu.N. Gornostyrev, D.L. Novikov, O.N. Mryasov,and A.J. Freeman, unpublished.
10. A.J. Duncan, M.J. Kaufman, C.T. Liu, and M.K. Miller, Appl.Surf.Sci, **76/77**, p. 155 (1994).
11. T. Hong and A.J. Freeman, Phys.Rev.B, **43**, p. 6446 (1991).
12. C.L. Fu and M.H. Yoo in High temperature Ordered Intermetallic Alloys IV, edited by L. Johnson, D.P. Pope and J.O. Stiegler (Mater. Res. Soc. Proc. 123, Pittsburg, PA 1991), p. 667.
13. M.J. Mills and D.B. Miracle, Acta Met. Mater. **41**, p. 85 (1993).
14. A.O. Anokhin, M.L. Galperin, Yu.N. Gornostyrev, M.I. Katsnelson and A.V. Trefilov, Pisma Zh. Eksp. Theor. Fiz., **59**, p. 198 (1994); Phys. Metall. Metallography, **79**, p. 16 (1995).

COHESIVE ENERGIES OF Be AND Mg CHALCOGENIDES

M. PORCU, G. SATTA, F. CASULA AND G. MULA
INFM and Dipartimento di Scienze Fisiche dell'Universita' di Cagliari
I - 09124 Cagliari - ITALY

ABSTRACT

We have calculated the cohesive energies, bulk moduli and equilibrium volumes of Be and Mg oxides, sulphides and selenides, in both zincblende and rocksalt structures. The calculations have been performed with the Discrete-Variational-Method (DVM), a real space first-principle local-density-functional approach. Comparisons with the experiment and with other first-principles approaches show that the electronic and structural properties of solids can be computed with DVM at least as accurately as with the usual plane-wave pseudopotential methods. This result is especially interesting in view of the fact that an order N implementation of DVM, based on the W. Yang's divide and conquer method, has been recently developed.

INTRODUCTION

First-principle pseudopotential methods[1] within the local density-approximation[2] (LDA) have abundantly proved their value in the calculation of total energy and related properties in almost every kind of solids, irrespective of the bonding and of the conduction type. Though in principle resorting to real space techniques is at least as good a choice, in practice only recently the emergence of order N methods[3] to tackle large scale computational problems has brought again to the forefront of research the use of localized orbitals. A need therefore exists for a reappraisal of real space methods in total energy calculations, and for a thorough testing of their accuracy and reliability. In this paper we offer a first evaluation of the Discrete Variational Method[4] (DVM) in the study of the binding energies of beryllium and magnesium chalcogenides.

Beryllium and magnesium chalcogenides are in fact wide band-gap semiconductors of technological and scientific interest because of their hardness and of their potential applications for blue-green laser diodes. In addition their joint study offers a particularly interesting challenge for a single theoretical approach since BeS and BeSe crystallize in the cubic zincblende structure while, in spite of their close chemical similarity, MgS and MgSe crystallize in the rocksalt structure. Furthermore BeO ground state structure is wurtzite while MgO one is rocksalt. The high pressure behaviour of these compounds is also a suitable testing ground for phase stability investigations, since e.g. differences by more than a factor of 6 in LDA pseudopotential predictions of transition pressures have been found for BeO[5] and further checkings with alternative methods have been called for.

A further motivation to study these compounds with DVM is the fact that all of them have been recently thoroughly investigated with a variety of methods, most of them first-principles ones. An impressive set of theoretical results, which range from the equilibrium lattice constants to bulk moduli, elastic constants and binding energies in zincblende, wurtzite and rocksalt phases and include estimates of the transition pressures, is thus available for comparison. While the experimental evidence is not so abundant, part of these results are in good agreement with the experiment, part are merely predictions. However, since our main goal is to test the reliability of a real space method such as DVM against the other available theoretical

tools, even comparisons of theoretical predictions obtained by different methods can be fully meaningful.

In this preliminary report we shall confine ourselves to zincblende and rocksalt structures, and to the calculation of cohesive energies, bulk moduli and equilibrium volumes. Moreover all calculations have been performed with a minimal basis set plus an extra set of p orbitals. Further properties and crystal structures, together with the extension to BeTe and MgTe phases, will be taken into account in a forthcoming paper[6], where an optimization of the basis will be carried out.

THE DISCRETE VARIATIONAL METHOD

DVM has been originally developed in the seventies by Ellis and Painter[7] to perform band structure calculations, its basic idea being to capitalize on the high efficiency of a numerical representation of localized orbitals to make their use competitive against plane waves expansions. In the case of solid state calculations this attempt has been effectively frustrated by the advent of norm conserving pseudopotentials[8], and DVM has since evolved into a full-fledged first-principles method specialized to deal with clusters[4,9]. We have undertaken the job of taking one of the latest versions[10] and converting it to an object-oriented architecture by embedding most of its Fortran instructions into a C++ framework[11]. In this way we have been able to retain the optimisation of the key numerical routines of the original code while enjoying the flexibility of the object-oriented C++ language. The advantages of this new architecture has greatly eased the subsequent transitions, again to solid state calculations and to W. Yang's order N approach[12].

Our computational procedure is thus based on a DVM code which straightforwardly implements all-electron density-functional calculations within LDA. We have used the Von Barth-Hedin formulation of the exchange and correlation contribution. The wavefunctions are expanded into a set of orbitals, obtained by solving numerically the isolated atom problem, and a grid of 2000 points/atom is used in numerical integrations in the unit cell. The use of such a high density grid has been made necessary by the convenience[6] of using the highly efficient diophantine algorithm even in the framework of integrations over unit cells (instead of the whole space). We have used the frozen core approximation to avoid spurious energy minimization effects on our structural properties. The cohesive energy has been computed for sets of molecular volumes for both structures of each compond and the results have been fitted to the Murnaghan equation of state.

RESULTS AND DISCUSSION

Our cohesive energies per atom versus the atomic volume, for zincblende and rocksalt structures, are shown in Figs. 1(a)-1(c) and in Figs. 1(d)-1(f) for Be and Mg chalcogenides, respectively. The curves predict that of the two structures zincblende is the most stable one for Be compounds, while the reverse should happen for Mg compounds. Both predictions are in agreement with the experiment, though the experimental stable structure of BeO is wurtzite, which we did not take into account, but which has the same coordination number as zincblende and usually gives rise to a cohesive energy curve which runs parallel and quite close to the zincblende one.
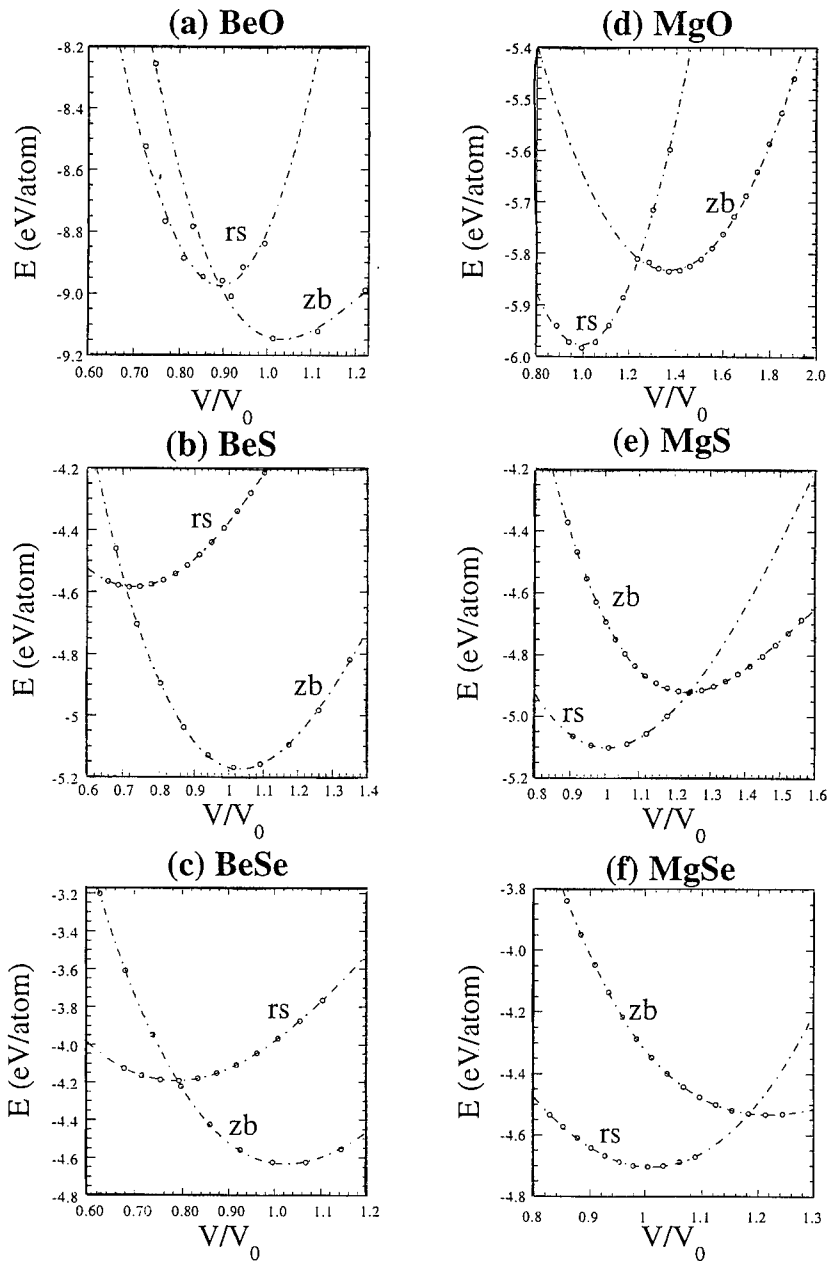
FIG.1. Cohesive energies of Be and Mg chalcogenides vs relative volume V/V_0 are shown in plots a-c and d-f respectively . Rocksalt (rs) and zincblende (zb) structures are considered. Calculated values and fitting curves are shown as well.

Another feature that is apparent from the curves is the shift in the relative positions of the equilibrium atomic volumes for the two structures, which means that in the case of Be chalcogenides one can, by the application of a suitable pressure, transform the zincblende stable structure into a rocksalt one, while this transformation is impossible for the analogous Mg compounds. For the latter a transition under pressure to a higher coordination structure such as cesium chloride should still be possible and seems indeed to be verified [13].

More quantitative comparisons are allowed by tables I - III. While our cohesive energies show the usual overestimation with respect to the experiment that is typical of local-density calculations, the estimates of the energy differences between the two structures taken into account are in line with the other theoretical calculations.

Table I - Cohesive energies E_c, lattice constants a_0, and bulk moduli B for zincblende and rocksalt structures of BeO, BeS and BeSe, obtained with the DVM method and compared with the available experimental data and other recent theoretical values.

	BeO		BeS		BeSe	
	rocksalt	zincblende	rocksalt	zincblende	rocksalt	zincblende
E_c(eV/atom)						
DVM	8.9	9.1	4.6	5.2	4.2	4.6
expt.						
others	6.61[1]	7.02[1]		$(\Delta=0.6)^2$		$(\Delta=0.6)^2$
a_0(Å)						
DVM	3.65	3.839	4.41	4.90	4.75	5.183
expt.		3.797[3]		4.865[5]		5.139[5]
others	3.589[1]	3.767[1]	4.492[4]	4.745[4]	4.795[4]	5.037[4]
B(GPa)						
DVM	307.6	260	101.6	86.9	104.4	86.7
expt.		212[6]				92.2[7]
others	272[1]	229[1]	127.5[4]	116[4]	91.2[4]	98.8[4]

1. FPLMTO results from J.C. Boettger and J.M. Wills, Phys. Rev. B54, 8965 (1996)
2. $\Delta = E_c$(zincblende) - E_c(rocksalt) from 4)
3. A. Lichanot and M. Rerat, Chem. Phys. Lett. 211, 249 (1993)
4. A. Muñoz, P. Rodríguez-Hernández and A. Mujica, Phys. Rev. B 54, 11861 (1996)
5. Data in Science and Technology. Semiconductors: Others than Group IV elements and III-V Compounds, edited by O. Madelung (Springer-Verlag, Berlin, 1992)
6. Wurtzite phase data from R.M. Hazen and L.W. Finger, J. Appl. Phys. 59, 3728 (1986).
7. H. Luo, K. Gandehari, R.G. Greene, A.L. Ruoff, S.S. Trail and F.J. Di Salvo, Phys. Rev. B52, 7058 (1995)

Inspection of the tables also shows that DVM equilibrium lattice constants compare quite well with the experiment for the experimental stable structure and with previous theoretical results for the other one. It is to be noted that for the stable structure our values are closer to the experiment that pseudopotential based methods. For BeO, where accurate all-electron calculations are available, our result has an absolute error quite close to the one of a full-potential linear muffin-tin orbital (FPLMTO) method[5]. We are also on the mark with zincblende MgS, for which an experimental estimate of the lattice constant has been possible from the study of zincblende ZnSe/MgS superlattices[14].

TABLE II - Cohesive energies E_c, lattice constants a_0, and bulk moduli B for zincblende and rocksalt structures of MgO, MgS and MgSe, obtained with the DVM method and compared with the available experimental data and other recent theoretical values.

	MgO		MgS		MgSe	
	rocksalt	zincblende	rocksalt	zincblende	rocksalt	zincblende
E_c(eV/atom)						
DVM	5.9	5.8	5.1	4.9	4.7	4.5
expt.	5.15[1]		4.0[1]		3.4[1]	
others	4.85[2]		4.50[3]	4.49[3]	4.08[4]	4.04[4]
a_0(Å)						
DVM	4.05	4.7	5.21	5.59	5.476	5.81
expt.	4.21[1]		5.20[6]	5.59[7]	5.463[8]	5.89[9]
others	4.27[2]		5.14[3]	5.59[3]	5.50[4]	5.98[4]
B(GPa)						
DVM	153.7	126.2	90.3	68.3	89.6	62.3
expt.	160[10]					
others	150[2]	151[11]	77.7[3]	57[3]	65.4[4]	47.8[4]

1. From the CRC Handbook of Chemistry and Physics, 70th edition.
2. P. Cortona and A. Villafiorita Monteleone, J. Phys.: Condens. Matter **8**, 8983 (1996)
3. Sun-Ghil Lee and K.J. Chang, Phys. Rev. B**52**, 1918 (1995)
4. P.E. Van Camp, V.E. Van Doren and J.L. Martins, Phys. Rev. B**55**, 775 (1997)
5. *Data in Science and Technology*. Semiconductors:Others than Group IV elements and III-V Compounds, edited by O. Madelung (Springer-Verlag, Berlin, 1992)
6. R.W.G. Wyckoff, *Crystal Structures* (Wiley, New York, 1963)
7. K. Uesugi, T. Obinata and I. Suemune, Appl. Phys. Lett. **68**, 844 (1996)
8. *Inorganic Compounds*, Crystal Data, Vol. II, edited by J.H. Donnay and H.M. Ondik (U.S. Department of Commerce, Washington DC, 1972)
9. H. Okuyama, K. Nakano, T. Miyajima and K. Akimoto, J. Cryst. Growth **117**, 139 (1992)
10. Z.P. Chang and E.K. Graham, J. Phys. Chem. Solids **38**, 1355 (1977)
11. N.C. Pyper, Phil. Trans. R. Soc. A**352**, 89 (1995)

Table III - DVM values of transition pressures P_T from zincblende to rocksalt structure and of corresponding transition volumes (normalized to zincblende equilibrium volumes) for BeO, BeS and BeSe, compared with other recent theoretical values.

	BeO		BeS		BeSe	
	zincblende	rocksalt	zincblende	rocksalt	zincblende	rocksalt
P_T(GPa)						
DVM	31.3		35.8		23.4	
others	94.1[1] , 21.7[2]		52.35[3]		42.76[3]	
V_T/V_0						
DVM	0.908	0.784	0.811	0.544	0.845	0.634
others	0.776[1], 0.930[2]	0.688[1], 0.742[2]	0.757[3]	0.651[3]	0.759[3]	0.659[3]

1. FPLMTO results from J.C. Boettger and J.M. Wills, Phys. Rev. B**54**, 8965 (1996)
2. *Ab initio* pseudopotential results for the wurtzite-rocksalt transition from K.J. Chang, and M.L. Cohen: Sol. St.Commun. **50**, 487 (1984)
3. *Ab initio* pseudopotential results for the wurtzite-NiAs transition from A. Muñoz, P. Rodríguez-Hernández and A. Mujica, Phys. Rev. B **54**, 11861 (1996)

As for our calculated bulk moduli the tables show that they follow the same trends and have the same kind of agreement with the experiment as the pseudopotential calculations by Lee and Chang[15], Muñoz et al. [16, 17] or Van Camp et al.[18], or the all-electron FPLMTO ones by Boettger and Wills[5]. As is well known, of course, agreement with the experiment is significantly worse than with the equilibrium lattice constants.

DVM predictions for transition pressures, reported in table III for Be chalcogenides (no pressure-induced transition to zincblende structure being feasible for Mg chalcogenides), confirm the general agreement with previous LDA results for the transitions between the same or very similar pair of structures. However in the case of BeO the spread of numerical predictions strongly suggests further more accurate investigations, both theoretical and experimental.

CONCLUSIONS

A systematic investigation of the structural properties of the chalcogenides of Be and Mg has been made with a real space method like DVM. In spite of the use of a not optimized basis set our results clearly show the reliability of the method and its applicability over a wide range of bonding types, from elemental semiconductors (studied in Ref. 10) to strongly ionic compounds such as BeO. On this basis its usefulness as an investigation means, both on its own and as the heart of an order N computational engine, can be stated with confidence.

REFERENCES

1. e.g. M.T. Yin and M.L. Cohen, Phys. Rev. B26, 5668 (1982).
2. For a review on LDA calculations see, e.g., R.O Jones and O. Gunnarson, Rev. Mod. Phys. 61, 689 (1989).
3. W.Yang, Phys. Rev. Lett. 66, 1438 (1991). See also ref. 12 and references therein.
4. J. Yang, F. Toigo and W. Kelin, Phys. Rev. B50, 7915 (1994) and references therein.
5. J.C. Boettger and J.M. Wills, Phys. Rev. B 54, 8965 (1996).
6. M. Porcu, G. Satta, F. Casula and G. Mula, to be published
7. D.E. Ellis and G.S. Painter, Phys. Rev. B2, 2887 (1970).
8. D.R. Hamann, M. Schlüter and C. Chiang, Phys. Rev. 43, 1494 (1979).
9. J. Yang, W. Kelin, F. Casula and G. Mula, Phys. Rev. B47, 4025 (1993).
10. B. Delley, D.E. Ellis, A.J. Freeman, E.J. Baerends and D. Post, Phys. Rev. B27, 2132 (1983).
11. G.Mula, C. Angius, F. Casula, G. Maxia, M. Porcu and J. Yang, Mat. Res. Soc. Proc. 408, 3 (1996).
12. W. Yang and T. Lee, J. Chem. Phys. 163, 5674 (1993), and W. Yang, Phys. Rev. B56, 9294 (1997).
13. As quoted in P.E. Van Camp, V.E. Van Doren and J.L. Martins, Phys. Rev. B55, 775 (1997) For MgO, however, no phase transition has been found up to pressures of 227 GPa: see T.S. Duffy, R.J. Hemley, and H. Mao, Phys Rev. Lett. 74, 1371 (1995).
14. K. Uesugi, T. Obinata and I. Suemune, Appl. Phys. Lett. 68, 844 (1996).
15. Sun-Ghil Lee and K.J. Chang, Phys. Rev. B 52, 1918 (1995).
16. A. Muñoz, P. Rodríguez-Hernández and A. Mujica, Phys. Rev. B 54, 11861 (1996).
17. M. Gonzáles-Díaz, P. Rodríguez-Hernández and A. Muñoz, Phys. Rev. B 55, 14043 (1997).
18. P.E. Van Camp, V.E. Van Doren and J.L. Martins, Phys. Rev. B55, 775 (1997)

QUANTUM MONTE CARLO SIMULATIONS OF DISORDERED MAGNETIC AND SUPERCONDUCTING MATERIALS

R.T. SCALETTAR[1], P.J.H. DENTENEER[2], C. HUSCROFT[1], A. MCMAHAN[3],
R. POLLOCK[3], M. RANDERIA[4], N. TRIVEDI[4], M. ULMKE[5], G.T. ZIMANYI[1]
[1]Physics Department, University of California, Davis, CA 95616
[2]Lorentz Institute, Leiden University, 2333 CA Leiden, The Netherlands
[3]Lawrence Livermore National Laboratory P. O. Box 808, Livermore, CA 94551
[4]Theor. Phys. Group, Tata Institute of Fundamental Research, Bombay 400005, India
[5]Theor. Phys. III, Institut für Physik, Universität Augsburg, D–86135 Augsburg, Germany

ABSTRACT

Over the last decade, Quantum Monte Carlo (QMC) calculations for tight binding Hamiltonians like the Hubbard and Anderson lattice models have made the transition from addressing abstract issues concerning the effects of electron-electron correlations on magnetic and metal-insulator transitions, to concrete contact with experiment. This paper presents results of applications of "determinant" QMC to systems with disorder such as the conductivity of thin metallic films, the behavior of the magnetic susceptibility in doped semiconductors, and Zn doped cuprate superconductors. Finally, preliminary attempts to model the Kondo volume collapse in rare earth materials are discussed.

INTRODUCTION

The determinant QMC method[1] is a powerful technique for understanding the physics of itinerant, interacting electrons. Its primary strength is that it treats the correlations between electrons exactly, in contrast to other approaches which resort to various simplifying approximations. The chief disadvantage is its computational cost, which limits the complexity of the models which can be considered. Many of the past applications[2] have been to the single–band two–dimensional Hubbard Hamiltonian. This model is of theoretical interest since it is the simplest lattice Hamiltonian exhibiting both an interaction–driven ("Mott") metal–insulator transition and also long range magnetic order. It is also potentially of importance in understanding the magnetic and superconducting properties of the CuO_2 sheets of high temperature superconductors. That the model has a single band and is in two–dimensions has played a crucial role in making simulations on reasonable lattice sizes (up to 16x16 sites) possible.

As algorithms and machine speeds have improved, however, computational restrictions are becoming less prohibitive, and the determinant QMC approach is being applied to tight–binding models which include features such as many orbitals, disorder, and higher dimensionality. In this paper we will provide an overview of four such applications. First, we have studied a model of disordered superconducting films, where there has been a long–standing interest in the possibility of a universal conductivity at the superconductor-insulator phase transition.[3] Second, we examine the effect of topological randomness on magnetically ordered phases, where one issue is the enhancement of the uniform spin susceptibility at low temperatures observed in doped semiconductors.[4] We next describe the behavior of magnetic correlations when non–magnetic impurity sites are introduced, a problem under investigation with recent experiments on Zn doping of ladder compounds

and high temperature superconductors.[5] Finally, we describe preliminary results on the phase diagram of the periodic Anderson Hamiltonian in three dimensions.[6] Here the key question is constructing a minimal model which might contain the essential features necessary for describing the "volume-collapse" transition observed in many rare earth systems.

In the remainder of this introduction, we will describe in more detail the physics of these problems. Following this, section 2 contains a discussion of the QMC algorithm. Sections 3, 4, and 5 contain the results of our simulations, and section 6 is a brief conclusion.

i. Most bulk metals, and many alloys, undergo a superconducting phase transition in which the resistivity plunges abruptly to zero as the temperature is lowered. Over the last two decades, this transition has been exhaustively studied for thin films.[7] When small amounts of Pb or Sn, for example, are deposited on a substrate, the resistance increases as the temperature is lowered; the material is an insulator. However, as more material is deposited, the resistance goes to zero as the temperature is lowered. That is, there is a superconducting-insulator transition controlled by the film thickness. These experiments have been carried out for a large variety of materials and substrates, and were also performed with different tuning parameters for the transition, for example changing the magnetic field strength in Bi or oxygen concentration in $In_{1-x}O_x$ samples at fixed thickness.

The remarkable observation is that the "critical resistance" which separates systems which are insulating from those which are superconducting is nearly the same for all materials. This is, of course, surprising, since such universality is normally expected in quantities like critical exponents, not in the coupling strength determining the location of a transitions. There have been a number of analytic attempts to explain this phenomenon,[8] but until very recently numeric approaches have been restricted to simplified "boson" models in which the electrons are treated as preformed Cooper pairs.[8, 9]

In section 3 we describe QMC calculations for a model of interacting, two-dimensional electrons which is known to display a superconducting phase transition as the temperature is lowered. We add disorder to this model, and show that a superconductor-insulator phase transition is induced, and evaluate the resistance at the transition.

ii. Doped semiconductors like Si:P exhibit a number of interesting magnetic phenomena. The basic physics is that of a two-fluid system with both itinerant electrons and randomly positioned local moments.[10] One essential question concerns whether there is a low temperature divergence of the magnetic susceptibility $\chi(q = 0)$. It has been suggested that random exchange couplings between moments can lead to isolated spins which have a Curie-like contribution to χ. Similar rapid increases in the susceptibility as the temperature is decreased are observed in the quasi-1D TCNQ compounds,[11] and are likewise believed to have their origin in randomness. Section 4a contains a discussion of simulations of a model of itinerant interacting electrons with hopping disorder which captures this behavior.

Another magnetic system in which controlled disorder can be studied is Zn substitution in ladder compounds[12] and in the cuprate superconductors.[13] There it is observed that antiferromagnetic order can be stabilized by randomness. Recent numerical studies have shown that, indeed, short range magnetic correlations can be enhanced by non-magnetic defects.[14] Here we present, in section 4b, a discussion of the effect of non-magnetic impurities on *long range* magnetic correlations and on the Mott metal-insulator transition.

iii. Many rare earth systems undergo a transition in which the volume is abruptly and drastically reduced as the pressure increases. For example, Ce, a particularly simple

case because the transition is isostructural, exhibits a first order phase transition with a 15% change in volume at 8 kbar. There are a number of competing theories for the origin of this effect.[15, 16] At the crudest level, a transfer of electrons between localized f and itinerant d orbitals might account for the volume change. Many theories assume that electron correlations play a crucial role, because the transition involves a loss of magnetic moments.

In section 5, we describe QMC simulations of the three-dimensional periodic Anderson model (PAM). This Hamiltonian has the requisite complexity to allow for the processes which may underly the volume collapse transition, namely separate conduction and localized orbitals and the potential for charge transfer between them; and also the possibility of "singlet" formation between electron spins in the two orbitals which would account for the disappearance of magnetic moments.

THE DETERMINANT QUANTUM MONTE CARLO METHOD

Determinant QMC has proven very useful for evaluating correlation functions in tight–binding problems of interacting electrons. Consider the Hubbard Hamiltonian,

$$\hat{H} = -\sum_{\langle \mathbf{i},\mathbf{j}\rangle\sigma} t_{\mathbf{i}\mathbf{j}} \left(c_{\mathbf{i}\sigma}^\dagger c_{\mathbf{j}\sigma} + c_{\mathbf{j}\sigma}^\dagger c_{\mathbf{i}\sigma} \right) + \sum_{\mathbf{i}} \left[U_{\mathbf{i}}(n_{\mathbf{i}\uparrow} - \frac{1}{2})(n_{\mathbf{i}\downarrow} - \frac{1}{2}) - \mu_{\mathbf{i}}(n_{\mathbf{i}\uparrow} + n_{\mathbf{i}\downarrow}) \right], \qquad (1)$$

which consists of a kinetic energy term describing the hopping of electrons between nearest–neighbor sites $\mathbf{i}$ and $\mathbf{j}$, an interaction term giving a repulsion between electrons of opposite spin species on the same site, and a chemical potential term which controls the filling.[17] Since our interest will be in disordered models, we have explicitly indicated that the transfer integrals $t_{\mathbf{i}\mathbf{j}}$, interaction strengths $U_{\mathbf{i}}$, and local chemical potentials $\mu_{\mathbf{i}}$ may be random. The Hamiltonian Eq. 1 is "particle–hole symmetric" at $\mu_{\mathbf{i}} = 0$ so that each site is precisely "half–filled", $\langle n_i \rangle = 1$, regardless of the values of the hopping and interactions (in the absence of charge–density–wave formation).

In general, QMC methods map a quantum problem in d spatial dimensions to a classical problem in $d + 1$ dimensions, where the inverse temperature $\beta = 1/T$ gives the additional length to the lattice. The reason is that in a classical statistical mechanics problem, the Boltzmann factor, $\exp[-\beta E]$, is a real–valued function of the degrees of freedom. In a quantum mechanical problem the relevant quantity is $\exp[-\beta \hat{H}]$ where $\hat{H}$ is an *operator*.

This exponential of the Hamiltonian operator cannot be computed without diagonalizing $\hat{H}$ (that is, solving the problem exactly). However, it is often the case that the exponential of the kinetic and potential energy pieces of $\hat{H}$ can be separately diagonalized, for example by working in momentum– or real– space respectively. Unfortunately, these constituent pieces do not commute, so the full exponential cannot be broken up into the product of the individual exponentials. The trick is to discretize the inverse temperature β into a large number L of much smaller divisions $\Delta\tau = \beta/L$ and use the "Trotter approximation,"[18]

$$Z = \text{Tr}[e^{-\beta\hat{H}}] = \text{Tr}[e^{-\Delta\tau\hat{H}}]^L \approx \text{Tr}[e^{-\Delta\tau\hat{K}} e^{-\Delta\tau\hat{V}}]^L. \qquad (2)$$

The discretization of β has introduced a small parameter and breaking up the exponential is now a procedure which becomes arbitrarily accurate as L increases. This discretization of β generates the added "inverse temperature" dimension to quantum simulations. In the

case of the Hubbard Hamiltonian, $\hat{K}$ includes the one body terms, the kinetic energy and chemical potential, and $\hat{V}$ includes the two–body interactions.

In the "world–line" algorithm, complete sets of states are introduced into Eq. 2, and the resulting matrix elements are evaluated analytically. The determinant QMC algorithm instead reduces the exponential of $\hat{V}$ to an exponential of a one body operator by introducing an integration over an auxiliary "Hubbard–Stratonovich" field,

$$e^{-\Delta\tau U\left(n_{i\uparrow}-\frac{1}{2}\right)\left(n_{i\downarrow}-\frac{1}{2}\right)} = \frac{1}{2}e^{-\Delta\tau U/4}\sum_{S(i,l)=\pm 1} e^{-\lambda S(i,l)\left(n_{i\uparrow}-n_{i\downarrow}\right)}. \tag{3}$$

The coupling constant λ is given by $\cosh(\lambda) = \exp[\Delta\tau U/2]$. We must introduce a field variable $S(i,l)$ at each spatial site i of the lattice where an interaction is present, and also at each imaginary time slice l in the Trotter decomposition of $\exp[-\beta\hat{H}]$.

The final step is to evaluate the trace over the fermion degrees of freedom, which can now be performed analytically since the exponentials involve only one-body terms.

$$Z = \sum_{\{S(i,l)\}=\pm 1} \det M_\uparrow(S)\det M_\downarrow(S). \tag{4}$$

Detailed forms for the matrices M_σ, which have dimension the number of spatial sites in the lattice, are given in the literature.[1] To summarize Eq. 4, the quantum mechanical partition function has been expressed in terms of an equivalent classical problem in one higher dimension; Z is a sum over a classical field indexed by i and l with an argument which no longer involves operators but is instead an effective classical Boltzmann weight.

We conclude with some general comments. First, the scaling of the algorithm is as the cube of the number of sites,[19] because of computations of the determinants of the matrices M_σ. Simulations of about 100 sites are already fairly computationally intensive, indicating why most applications have been to the simplest single orbital models in low dimension. Second, we are actually interested in measuring operator expectation values, $\langle \hat{A} \rangle = Z^{-1}\mathrm{Tr}[\hat{A}e^{-\beta\hat{H}}]$. Here, $\hat{A}$ might be the charge, spin, superconducting pair, or current density on a site or a "correlation function" consisting of a product of two such densities on different sites of the lattice. Such traces can easily be formulated in the same way as outlined above for Z, and, indeed, measurements of such physical quantities involve only accumulating the appropriate elements of the inverse of M_σ. Finally, there is a very important limitation to determinant QMC simulations. The determinants can become negative, and in such cases measurements can have impractically large variance. In this paper we will deal with situations where this problem does not occur.[20]

The calculations presented here will be mainly results from determinant QMC, but we will also show a few computations using "dynamical mean–field theory" (DMFT).[21] This technique also employs an auxiliary–field QMC algorithm, but introduces a local approximation to the self-energy, which becomes exact in the limit of infinite dimensions, and allows a more complete exploration of parameter space.

DISORDER–DRIVEN SUPERCONDUCTOR–INSULATOR PHASE TRANSITIONS

We first study a Hamiltonian of the form given in Eq. 1 with uniform hoppings $t_{ij} = t$ and uniform *attractive* interactions $U_i = -|U|$. In the absence of randomness in the chemical potential, this "attractive Hubbard model" is known to exhibit a Kosterlitz–Thouless phase transition to a superconducting state as the temperature is lowered off half-filling.
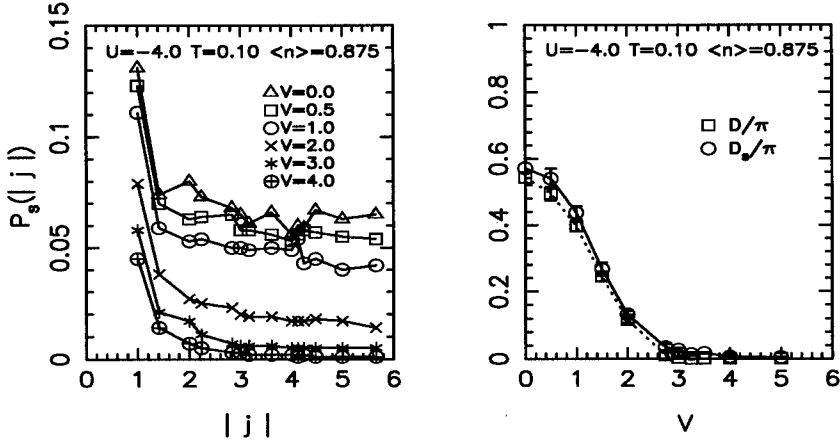
Figure 1: Left panel: Pairing correlations as a function of lattice separation for different disorder strengths, $-V/2 < \mu_i < V/2$. For $V \approx 3$ the correlations become very small at large distances. Right panel: Superfluid stiffness D_s as a function of disorder strength. Again, for $V \approx 3$ this measure of superconductivity vanishes. The Drude weight D, the coefficient of $\delta(\omega)$ in the ac conductivity, $\sigma(\omega)$, is also shown, and equals D_s for this system. The lattice size is 8x8.

In Fig. 1 we show what happens to two quantities which measure whether the system is in a superconducting state. The pair correlation function,

$$p_s(\mathbf{j}) = \langle\, \Delta_{\mathbf{i+j}}\Delta_{\mathbf{i}}^{\dagger}\, \rangle \qquad\qquad \Delta_{\mathbf{j}}^{\dagger} = c_{\uparrow\mathbf{j}}^{\dagger}c_{\downarrow\mathbf{j}}^{\dagger}, \tag{5}$$

will go to a constant non–zero value at large separations $\mathbf{j}$ if the state is superconducting, and otherwise decay to zero.

The superfluid stiffness is a second quantity which is non–zero in the superconducting state. It is defined in terms of an appropriate limit of the current–current correlation function,

$$\begin{aligned}
D_s &= -K_x - \Lambda^{\mathrm{T}}, \\
K_x &= \langle -t\sum_{\sigma}(c_{\mathbf{1}+\hat{x},\sigma}^{\dagger}c_{\mathbf{1},\sigma} + c_{\mathbf{1},\sigma}^{\dagger}c_{\mathbf{1}+\hat{x},\sigma})\rangle, \\
\Lambda^{T} &\equiv \lim_{q_y\to 0}\ \Lambda_{xx}(q_x = 0, q_y; i\omega_n = 0), \\
\Lambda_{xx}(\mathbf{q}; i\omega_n) &= \sum_{\mathbf{1}}\int_0^{\beta} d\tau\langle j_x(\mathbf{1},\tau)j_x(0,0)\rangle e^{i\mathbf{q}\cdot\mathbf{1}}e^{-i\omega_n\tau}, \\
j_x(\mathbf{1},\tau) &= e^{H\tau}\left[it\sum_{\sigma}(c_{\mathbf{1}+\hat{x},\sigma}^{\dagger}c_{\mathbf{1},\sigma} - c_{\mathbf{1},\sigma}^{\dagger}c_{\mathbf{1}+\hat{x},\sigma})\right]e^{-H\tau}. \tag{6}
\end{aligned}$$

Here $\omega_n = 2n\pi/\beta$. In the right panel of Fig. 1 we show D_s as a function of the randomness V in the chemical potential. Disorder drives a transition to a state in which the pairing order is destroyed and the superfluid stiffness vanishes.
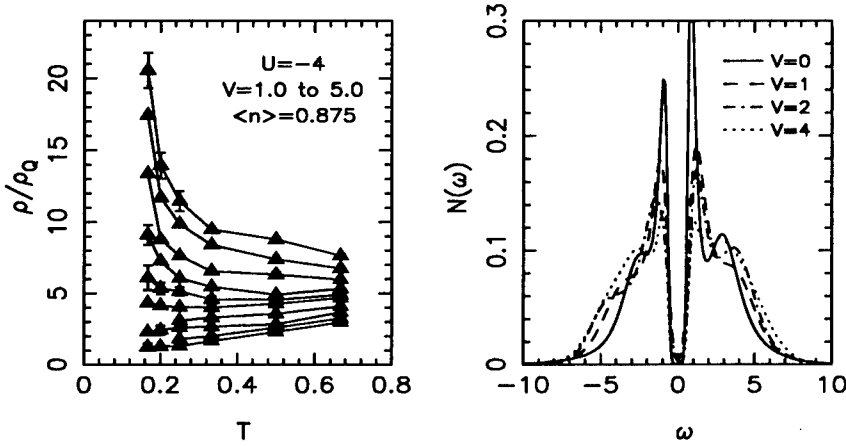
159

Figure 2: Left panel: dc resistivity as a function of temperature for different disorder strengths. For $V \approx 3$ the traces show a transition from superconducting ($d\rho/dT > 0$) to insulating ($d\rho/dT < 0$) behavior. $\rho_Q = \hbar^2/4e^2$ is the quantum of resistance. Right panel: Density of states $N(\omega)$ as we sweep through the transition. For both panels, the lattice size is 8x8.

What happens to the resistivity, ρ_{dc}, the experimentally most interesting quantity, as we cross through this transition? In the left panel of Fig. 2 we show ρ_{dc} as a function of temperature. This plot looks remarkably similar to that seen in experiments[7]: a family of traces exhibiting insulating behavior is separated from a family where the resistivity decreases as T is lowered, with a roughly constant separating curve at a critical disorder strength. (This value is consistent with estimates of the transition point obtained from the data of Fig. 1.) The value of the critical resistance is roughly the same as that obtained in previous numerical studies of bosonic models, which are the $|U|/t \to \infty$ limit of the Hamiltonian considered here.[8, 9] This value is about twice that observed experimentally, a disagreement which may be due to the lack of long range Coulomb repulsion in the model. We are currently studying this issue in greater detail.

The right panel of Fig. 2 shows the density of states (DOS). Interestingly, the gap in the density of states which exists in the superconducting phase is *not* destroyed when the disorder is increased to large enough values to suppress pair correlations and D_s (Fig. 1) to zero. This is in contrast to results[22] for the repulsive Hubbard model in which the analogous long range AF correlations appear to be necessary for the existence of a gap in $N(\omega)$. The DOS was obtained using the "maximum entropy algorithm."[23] Our results show that this technique works well even in the presence of disorder and the associated error bars from disorder averaging.

DISORDER–DRIVEN MAGNETIC PHASE TRANSITIONS

In the previous section we saw that random chemical potentials drive a superconducting–insulator transition and evaluated the resistivity at the transition, a quantity which exhibits

a roughly "universal" value experimentally. Just as disorder interferes with long range pairing order in the attractive Hubbard model, it also can affect magnetic correlations in the repulsive case. In this section we describe simulations of these effects in models which have been suggested as appropriate to P doped Si, and Zn doped $LaCuO_4$.

Random Bond Models

It is believed that the random hopping Hubbard Hamiltonian, and its strong coupling limit the random exchange Heisenberg Hamiltonian, might be appropriate models of the behavior of the magnetic susceptibility in doped semiconductors.[25] The idea is that as the randomness is increased, singlets can form on strong bonds, instead of the magnetically ordered phase in which all spins are correlated, which is the ground state in the clean limit. Depending on the topology, some single spins may be left isolated in this process of singlet formation, and these spins contribute a divergent Curie term to the magnetic susceptibility at low temperatures. While the Hubbard case has been studied analytically, and numerically on small lattices, QMC simulations in two dimensions have only recently been reported.[4] Here we review simulations of Eq. 1 in which the interactions $U_i = U > 0$ and chemical potentials $\mu_i = 0$ are constant, but the hopping is disordered, $t - \Delta/2 < t_{ij} < t + \Delta/2$. We show that indeed, anomalies in χ do appear.

The left panel of Fig. 3 shows the spin–spin correlation function,

$$C(1) = \langle M_{i+1}^z M_i^z \rangle = \langle (n_{i+1\uparrow} - n_{i+1\downarrow})(n_{i\uparrow} - n_{i\downarrow}) \rangle, \qquad (7)$$

as a function of separation for different disorder strengths. $C(1)$ exhibits a characteristic alternating pattern indicative of antiferromagnetism. Bond disorder steadily suppresses
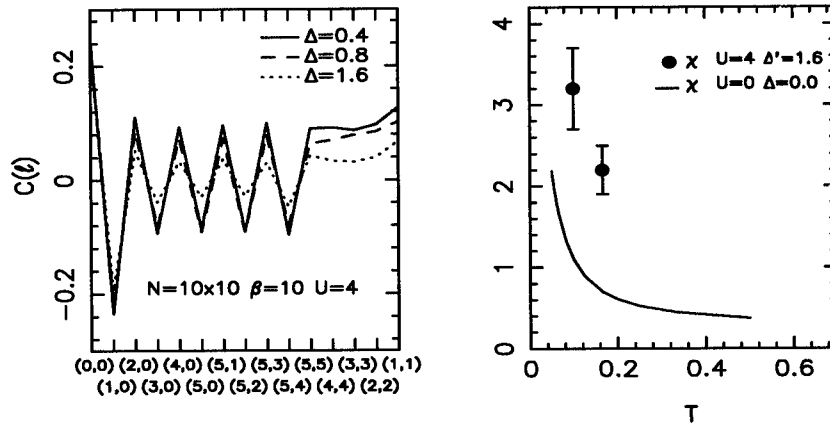

Figure 3: Left panel: Spin correlations as a function of lattice separation for different disorder strengths. For $\Delta \approx 1.6$, a finite size scaling analysis indicates the absence of long range order. Right panel: Uniform susceptibility as a function of temperature. In the presence of disorder *and* interaction, χ is enhanced. The "prime" on Δ indicates the use of a specific type of correlated disorder discussed in Ref. [4].

these correlations. A finite size scaling analysis indicates that for $\Delta > \Delta_c \approx 1.6$ the AF order parameter vanishes, and, instead, a singlet phase emerges.

The right panel of Fig. 3 exhibits the behavior of the uniform susceptibility,

$$\chi = \sum_1 \int_0^\beta d\tau \langle e^{\tau H} M_{i+1}^z e^{-\tau H} M_i^z \rangle, \tag{8}$$

for different parameter regimes. The solid curve shows the noninteracting limit. In the presence of interactions, and at a small disorder value, $\chi(0,0)$ is *reduced* below this $U = 0$ value. (The AF response, $\chi(\pi, \pi)$ is, of course, dramatically enhanced.) Meanwhile, for $\Delta \approx \Delta_c$ the uniform susceptibility instead shows a significant increase.

Nonmagnetic Impurity Models

We now turn to a model in which the hopping and chemical potential in Eq. 1 are uniform, but the interactions are random.[5] Specifically, we choose to set U_i to zero at some fraction f of the sites of the lattice. One motivation for studying such a model is a purely theoretical one. The Hubbard Hamiltonian at half–filling exhibits both a Mott metal–insulator transition when the energy cost U for double occupancy exceeds the bandwidth, and also AF order. There has recently been considerable discussion concerning whether a Mott gap can open in the absence of symmetry breaking such as that associated with AF order.[24] A model with a fraction f of sites with $U_i = 0$ might separate these two phenomena, since the Mott gap will be shifted from half–filling to a density $\langle n \rangle = 1 + f$, while AF order is likely to remain at half–filling due to commensuration effects.

An experimental motivation for the model is provided by studies of the effect of doping non–magnetic impurities like Zn into ladder compounds and oxide superconductors. One observation is that, in the latter case, AF order persists out to a Zn concentration of $x_c \approx 0.10 - 0.15$, far larger than the amount of doping with an isovalent element like Sr, $x_c \approx 0.03$, required to destroy AF.

In Fig. 4 we give results of DMFT simulations of the Hubbard Hamiltonian with $U_i = 0$ impurity sites. The left panel shows the phase diagram, and demonstrates the interesting feature that turning on randomness *enhances* the region of stability of the AF phase. In finite d the region of AF order at zero defect density, $f = 0$, will be shrunk down to half–filling $\langle n \rangle = 1$, but the enhancement remains and AF order persists significantly away from half–filling. Unfortunately, because of this stabilization of AF order out to the density $\langle n \rangle = 1+f$, we conclude this model does not exhibit a separation of Mott and AF behavior.

The right panel of Fig. 4 shows the DOS. At half–filling the DOS for the $U = 8t$ sites shows typical AF structure, two broad Hubbard bands, two quasi–particle peaks at low energies, and a gap at the Fermi energy. Surprisingly, there is an induced gap in the AF phase, even for the $U = 0$ sites, despite the absence of interactions there. The lower plot shows that at $n = 1 + f = 1.11$, the position of the Mott transition, there is still AF order.

VOLUME COLLAPSE TRANSITIONS

Several Lanthanides exhibit phase transitions under pressure characterized by abnormally large volume changes (14 % for Cerium and 9 % for Praseodymium). The physical mechanisms responsible for these transitions have been debated since discovery of the Cerium phenomenon over 50 years ago. Currently, the two main viable conjectures are a
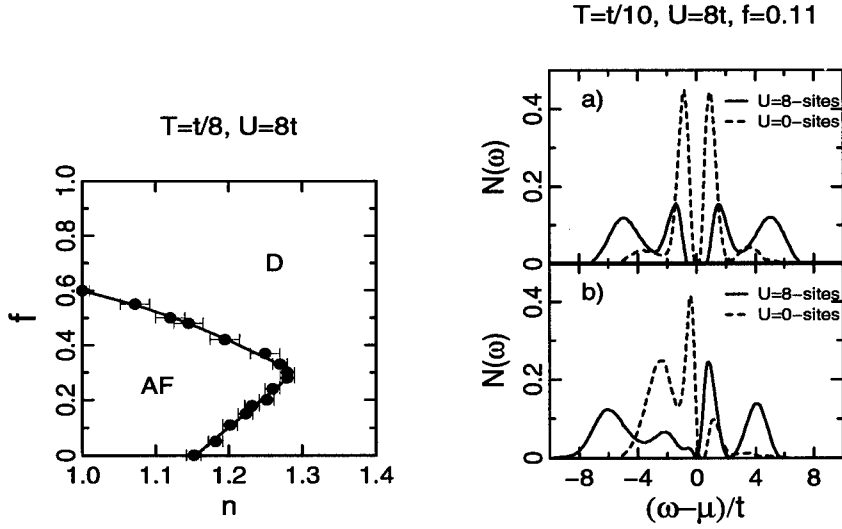
Figure 4: Left panel: Phase diagram of the Hubbard Hamiltonian as a function of filling and concentration of defect sites within dynamical mean field theory. The disorder extends the region of long range AF order to larger doping. A disordered phase (D) is dominant at high carrier and defect doping. Right panel: The density of states on the $U = 0$ and $U = 8$ sites at a) half-filling and b) $\langle n \rangle = 1.11$.

Mott transition of the 4f electrons accompanied by magnetic ordering[15] and a "Kondo volume collapse" due to rapid change in the 4f-valence electron coupling.[16] In both cases the strongly correlated nature of the 4f electrons is fundamentally implicated. To model such a transition it is necessary to consider systems with separate conduction and localized orbitals to allow for the transfer of charge between them and also for Kondo screening of the local moments by the conduction electrons. Indeed, our studies of single band models put the transition at grossly incorrect values of the ratio of interaction to bandwidth.[6]

In Fig. 5 we show the phase diagram we have obtained for the three dimensional periodic Anderson model. The precise Hamiltonian we have studied at this point consists of a cubic array of conduction sites hybridized to their near–neighbors by $t_{vv} = 1$, and also connected to localized orbitals ($t_{ff} = 0$), with an inter–orbital hybridization t_{fv} on near–neighbor sites. We have fixed the repulsion $U_{ff} = 8$ on the impurity sites, left the conduction electrons non–interacting, $U_{vv} = 0$, and explored the properties of the system as we vary temperature and t_{fv}.

We are still in the process of connecting these QMC studies to accurate calculations of the band structure in order to extract thermodynamic behavior and the details of the volume collapse transition.
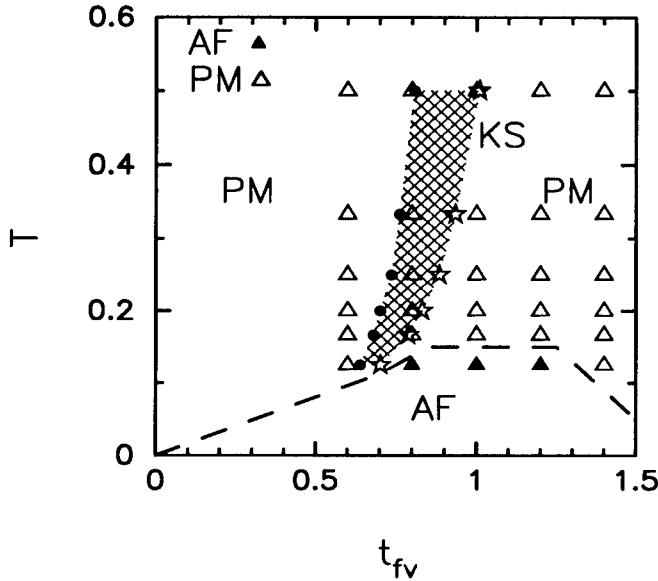
Figure 5: Phase diagram of the periodic Anderson model in three dimensions. At the lowest temperatures the local moments order antiferromagnetically (AF). At higher temperatures there are two distinct disordered regions, separated by a fairly narrow cross-over (hatched area). A paramagnetic (PM) region exists at small t_{fv} and a "Kondo singlet" (KS) region at large t_{fv}. These are distinguished by having small (large) values of correlation between near-neighbor f and d electron spins, respectively. The lattice sizes simulated were 4^3 and 6^3.

CONCLUSIONS

In this paper we have discussed a number of recent determinant QMC simulations of tight-binding Hamiltonians which incorporate some of the complicating features, including disorder, non-trivial orbital structure, and higher dimensionality, which are necessary to describe realistic materials. A significant amount of work remains to be done before these calculations provide compelling descriptions of the experimental properties of solids. While some problems have been solved, for example the extension of maximum entropy methods to obtain dynamical information in disordered systems, significant bottlenecks include adequate treatments of the sign problem to generalize the fillings which are accessible and also to make more practical the study of intersite and interorbital repulsion.

ACKNOWLEDGEMENTS

REFERENCES

1. R. Blankenbecler, D.J. Scalapino, and R.L. Sugar, Phys. Rev. **D24**, 2278 (1981).

2. D.J. Scalapino, Physics Reports **250**, 330 (1995); E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).

3. N. Trivedi, R.T. Scalettar, and M. Randeria, Phys. Rev. **B54**, 3756 (1996).

4. M. Ulmke and R.T. Scalettar, Phys. Rev. **B55**, 4149 (1997).

5. P. J. H. Denteneer, M. Ulmke, R. T. Scalettar, G. T. Zimanyi, Physica **A**, to appear. M. Ulmke, P. J. H. Denteneer, R. T. Scalettar, G. T. Zimanyi, unpublished.

6. C. Huscroft, A. McMahan, R.T. Scalettar, work in progress.

7. See review article by A. F. Hebard, in "Strongly Correlated Electronic Systems" edited by K. S. Bedell, Z. Wang, D. E. Meltzer, A. V. Balatsky, and E. Abrahams, (Addison-Wesley, 1994).

8. For a brief review, see M. Cha, M.P.A. Fisher, S.M. Girvin, M. Wallin, and A.P. Young, Phys. Rev. **B44**, 6883 (1991), and the references cited therein.

9. W. Krauth, N. Trivedi and D. M. Ceperley, Phys. Rev. Lett. **67**, 2307 (1991). E. S. Sorensen, M. Wallin, S. M. Girvin, and A. P. Young, Phys. Rev. Lett. **69**, 828 (1992). K.J. Runge, Phys. Rev. **B45**, 13136 (1992). G.G. Batrouni, B. Larson, R.T. Scalettar, J. Tobochnik, and J. Wang, Phys. Rev. **B48**, 9628 (1993).

10. For a recent review, see D. Belitz and T.R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).

11. L.N. Bulaevskii *et.al.*, Zh. Eksp. Teor. Fiz. **62**, p. 725 (1972) [Sov. Phys. JETP **35**, 384 (1972)]; and L.J. Azvedo and W.G. Clark, Phys. Rev. **B16**, 3252 (1977).

12. M. Azuma, Y. Fujishiro, M. Takano, T. Ishida, K. Okuda, M. Nohara, and H. Takagi (to be published); M. Nohara, H. Takagi, M. Azuma, Y. Fujishiro, and M. Takano (to be published).

13. G. Xiao et al., Phys. Rev. Lett. **60**, 1466 (1988); B. Keimer et al., Phys. Rev. B **45**, 7430 (1992); A. V. Mahajan, H. Alloul, G. Collin, and J. F. Marucco, Phys. Rev. Lett. **72**, 3100 (1994).

14. G.B. Martins, M. Laukamp, J. Riera, and E. Dagotto, Phys. Rev. Lett. **78**, 3563 (1997).

15. B. Johansson, Philos. Mag. **30**, 469 (1974).

16. J.W. Allen and R.M. Martin, Phys. Rev. Lett. **49**, 1106 (1982); and M. Lavagna, C. Lacroix, and M. Cyrot, Phys. Lett. **90A**, 210 (1982).

17. While we have used the term "site" for the indices **i** and **j**, these can in fact also label different orbitals. From a technical viewpoint, the determinant QMC algorithm treats distinct sites and orbitals in precisely the same way. They differ only in the Hamiltonian itself, that is in the choice of t_{ij}, U_i, and μ_i.

165

18. H.F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959); and M. Suzuki, Phys. Lett. **113A**, 299 (1985).

19. Naively the algorithm should scale as the fourth power of the number of sites N, since for each move of one of the NL Hubbard–Stratonovich fields, a calculation of the determinant requires N^3 operations. However, a trick reduces this by one power of N.

20. S.R. White, D.J. Scalapino, R.L. Sugar, E.Y. Loh, Jr., J.E. Gubernatis, and R.T. Scalettar, Phys. Rev. **B40**, 506 (1989).

21. W. Metzner and D. Vollhardt, Phys. Rev. Lett. **62**, 324 (1989); for a review see D. Vollhardt, in *Correlated Electron Systems*, edited by V. J. Emery (World Scientific, Singapore, 1993), p. 57; A. Georges, G. Kotliar, W. Krauth, and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996); M. Jarrell, Phys. Rev. Lett. **69**, 168 (1992); and T. Pruschke, M. Jarrell, and J. K. Freericks, Adv. Phys. **44**, 187 (1995).

22. M. Vekic and S.R. White, Phys. Rev. **B47**, 1160 (1993).

23. M. Jarrell and J. E. Gubernatis, Phys. Rep. **269**, 133 (1996).

24. F. Gebhard, *The Mott Metal–Insulator Transition*, Springer Tracts in Modern Physics, vol. **137** (Springer, Heidelberg, 1997).

25. M. Milovanovic, S. Sachdev, and R.N. Bhatt, Phys. Rev. Lett. **63**, 82 (1989); A.W. Sandvik, D.J. Scalapino, and P. Henelius, Phys. Rev. **B50**, 10474 (1994); and T.R. Kirkpatrick and D. Belitz, Phys. Rev. Lett. **76**, 2571 (1996).

MONTE CARLO STUDIES FOR STRONG CORRELATIONS IN HUBBARD-TYPE MODELS

E. S. Heeb
Institute for Theoretical Physics, ETH Zürich, CH-8093 Zürich, SWITZERLAND

ABSTRACT

The tight-binding model with repulsive Hubbard interactions represents an ideal prototype for the study of strong correlations. While exact numerical methods have been used with some success, they are typically limited by the size of the clusters that can be investigated or by the temperatures that can be reached. Variational methods, on the other hand, often require considerable advance knowledge of ground-state properties. The method presented here alleviates this problem by augmenting the variational approach with a scheme similar to Lanczos iterations thus bridging the gap between exact diagonalization and variational approaches. For the t-J model, the low-energy effective Hamiltonian of the Hubbard model, material properties like broken translational invariance or superconducting correlations are then investigated and a region of stability of a superconducting phase is found.

INTRODUCTION

The effects of strong correlations in condensed matter systems represent some of the most challenging problems in the description of materials. In the framework of the tight-binding approach the Hubbard model can be seen as the prototype for a system with strong correlations[1]. While originally it has been introduced in order to explain magnetic ordering the Hubbard model and numerous variations are now seen as an effective description for systems exhibiting a variety of different ground states. The discovery of high-temperature superconductors has renewed the interest in finding ground states with superconducting ordering in strongly correlated systems.

Despite considerable effort not many exact solutions for strongly correlated system have been found around which a perturbation expansion would be possible. So far our understanding of strong correlations is best for one-dimensional systems. This includes an exact solution for the canonical one-band Hubbard model[2] and a categorization of possible ground states by a renormalization approach[3]. Yet it is unclear whether or how these findings apply to higher dimensional systems, especially since one-dimensional systems cannot exhibit some of the relevant properties like long-range superconducting order in their ground state. In this situation numerical investigations have proven to be most valuable. Some of the most prominent are exact diagonalization of small clusters, a variety of quantum Monte Carlo simulations, and high-temperature series expansions. While these are some of the most powerful methods available, they have their limitations. Exact diagonalization is restricted to rather small systems and extrapolation to the thermodynamic limit is often questionable if not impossible. Quantum Monte Carlo methods suffer from the fermionic sign problem and temperatures close enough to the ground state cannot be reached before the results are swamped by statistical errors. Similarly, the number of terms in high-temperature expansions is the limiting factor in reaching low enough temperatures.

Variational methods have played a remarkably modest role in the investigation of strong correlations, while about at the same time they were crucial to the understand-

ing of the Quantum Hall effect. To be sure, also for these strongly correlated systems with unscreened Coulomb interactions there has been quite some effort to study variational approximations[4, 5], but so far no systematic scheme has emerged as it has been established with the Hartree-Fock and self-consistent field method for weak correlations. The variational results for strong correlations are usually seen as biased and remain controversial. This presentation tries to eliminate some of the bias by introducing systematic improvements to the variational wavefunctions. By proper analysis of the results we can achieve much of the confidence that is attributed to self-consistent field solutions.

THE METHOD

The variational method in quantum mechanics relies on the Rayleigh-Ritz principle. Given any variational wavefunction $|\Psi_{\text{var}}\rangle$ we can obtain an upper bound for the ground state energy of the Hamiltonian $\mathcal{H}$. The task is then to find an optimal wavefunction where the energy expectation value can be easily evaluated and which has the lowest upper bound. This method lends itself nicely to a Monte Carlo treatment[6]. If we choose a complete set of basis states $|\alpha\rangle$ we can write

$$E_{\text{GS}} \leq \frac{\langle \Psi_{\text{var}}|\mathcal{H}|\Psi_{\text{var}}\rangle}{\langle \Psi_{\text{var}}|\Psi_{\text{var}}\rangle} \tag{1}$$

$$= \sum_\alpha \frac{\langle \Psi_{\text{var}}|\alpha\rangle \langle \alpha|\mathcal{H}|\Psi_{\text{var}}\rangle}{\langle \Psi_{\text{var}}|\Psi_{\text{var}}\rangle} \tag{2}$$

$$= \sum_\alpha \frac{\langle \alpha|\mathcal{H}|\Psi_{\text{var}}\rangle}{\langle \alpha|\Psi_{\text{var}}\rangle} \frac{|\langle \alpha|\Psi_{\text{var}}\rangle|^2}{\langle \Psi_{\text{var}}|\Psi_{\text{var}}\rangle}. \tag{3}$$

This last equation has the form $F = \sum_x f(x)P(x)$. Taking $P(x)$ as a probability distribution we obtain a statistical estimate of F by sampling $f(x)$ according to the probability $P(x)$. The quantum mechanical probability has already the necessary conditions for a probability distribution so that no sign problem arises. Also, if $|\Psi_{\text{var}}\rangle$ is close enough to an eigenstate of the Hamiltonian then the sampled values for the energy do not scatter much and the convergence if the Monte Carlo sampling is quick. It should also be noted that the variance of the Hamiltonian can be calculated with only minimal extra effort.

$$\sigma_{\mathcal{H}}^2 = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 \quad \text{with} \quad \langle \mathcal{H}^2 \rangle = \sum_\alpha \left| \frac{\langle \alpha|\mathcal{H}|\Psi_{\text{var}}\rangle}{\langle \alpha|\Psi_{\text{var}}\rangle} \right|^2 \frac{|\langle \alpha|\Psi_{\text{var}}\rangle|^2}{\langle \Psi_{\text{var}}|\Psi_{\text{var}}\rangle}. \tag{4}$$

For weakly correlated systems the variational method is very successful. The ground state can very well be approximated by a Slater determinant of single electron orbitals. Taking the set of all such Slater determinants as our variational wavefunctions we arrive at the well known Hartree-Fock description of electrons. With a Bogolyubov transformation we can extend this same scheme to BCS wavefunctions which are used to describe superconductivity.

This scheme fails to work for strong correlations. As implied by the definition of strong correlations the electrons cannot be treated independently of each other and so a Hartree-Fock-Bogolyubov type of wavefunction is not any more a good approximation for the ground state. Most variational wavefunctions that are constructed to describe strongly correlated systems are somewhat ad hoc. They are usually regarded as biased and provide

only supportive evidence for the hypothesis that those wavefunctions are supposed to describe.

Nevertheless, the case for the variational method in investigating strong correlations is not as grim as it may seem at first. The reason why the independent electron approximation works for weak correlation is the observation that for each electron the others act as an effective background medium. While this is not true anymore for strong correlations, any long range correlations still remain and the orbitals get entangled only on the short scale where the strong correlations are most effective. Specifically, if we take superconducting long range order as an example we know that it is described by anomalous expectation values of the form $\langle c_i c_j \rangle$. We can describe a Cooper pair at position R with $\Delta(R) = \sum_{i,j} \varphi(i,j) c_{R+i} c_{R+j}$. Superconducting long range order is characterized by

$$\left\langle \Delta^\dagger(R)\Delta(0) \right\rangle \to \left\langle \Delta^\dagger(R) \right\rangle \left\langle \Delta\,(0) \right\rangle = \text{finite} \quad \text{for} \quad R \to \infty. \tag{5}$$

This is then evidence for a finite order parameter which in the case of superconductivity is described by the Ginzburg-Landau phenomenological theory.

For strongly correlated systems such an ordered phase is still connected to some long range order and the corresponding Hartree-Fock-Bogolyubov wavefunction would still describe the correct long range behavior. However it largely misses to represent the strong local correlations and therefore does poorly on the energy. We then need to improve the local correlations without changing the long range part of the wavefunction. It turns out that this can be done systematically.

A first approach to a systematic improvement of the local correlations are Lanczos iterations as they are used in exact diagonalization. Starting with a variational wavefunction $|\Psi_{var}^0\rangle$ the n-th Lanczos iteration will produce an improved wavefunction

$$|\Psi_{var}^n\rangle = C \left(1 + \mu_1 \mathcal{H} + \mu_2 \mathcal{H}^2 + \ldots + \mu_n \mathcal{H}^n \right) \left|\Psi_{var}^0\right\rangle, \tag{6}$$

where the μ_i are optimized to obtain the lowest energy and C is a normalization constant.

These Lanczos iterations lead to significant improvements in the energy. Test calculations for the 2-dimensional antiferromagnetic Heisenberg model[7] where the ground state energy is known from other numerical methods have shown that for some realistic starting wavefunctions improvements on the order of 80 % of the remaining correlation energy are possible in each Lanczos step. The calculations also show that the symmetry properties of the wavefunction have a great impact on the rate of convergence. It is more important that the starting wavefunction has the same symmetry properties as the ground state than that the energy expectation value for the starting wavefunction is lowered by including a number of correlations that happen to be easy to implement. This observation can be explained easily. Each Lanczos iteration suppresses higher excited states more and more. If we relax the requirements on the symmetry of the starting wavefunction we make it possible to mix in low-lying excited states with different symmetry properties. While this will lower the energy of the starting wavefunction it also makes it harder for the Lanczos iteration to project out the excited states. As a rule of thumb we can expect that if the starting wavefunction is composed of n energy eigenstates it will take $n-1$ Lanczos iterations to arrive at the ground state. Requiring the symmetry of the starting wavefunction to be the same what is known of the ground state is then a convenient way to speed up convergence even if it means to start from a slightly higher variational energy.

For practical applications it turns out that Lanczos iterations are not ideal. Using expectation values of the original wavefunction as we do in our approach, the time required

169

for the calculation increases exponentially with the number of Lanczos iterations. Already the second Lanczos iteration requires a great computational effort on any reasonably sized lattice and higher orders are possible only on small lattices. It is then advantageous to use the Lanczos method in the conventional way for exact diagonalization, where in each step the complete wavefunction is stored and the computational effort increases only linearly with the number of Lanczos iterations. Obviously we need to enhance our method. There is an important observation as to why the Lanczos iterations quickly become slow. While the first Lanczos iteration improves the local correlations due to the local nature of the Hamiltonian this is not entirely the case for the second and higher orders. The second application of the Hamiltonian on the variational wavefunction need not pick up where the first one left off. In this way the Lanczos iterations also affect the long range part of the wavefunction which in general they need to. However, in our scheme it is the task of an optimal choice of the variational parameters in the starting wavefunction which controls long range behavior. It is therefore a waste of computer time to include terms which act at long distances. We replace the Lanczos iterations with a generalized scheme using cluster operators

$$|\Psi_{\text{var}}^n\rangle = \sum_n \mu_n \mathcal{A}_n \left|\Psi_{\text{var}}^0\right\rangle, \tag{7}$$

where the local operators have the form

$$\mathcal{A}_n = \sum_g \mathcal{U}_g \mathcal{A}_n^{\text{local}} \mathcal{U}_g^{-1}. \tag{8}$$

Here all $\mathcal{A}_n^{\text{local}}$ are operators that act only on a given cluster and leave the rest of a configuration unchanged. The index g runs over all chosen symmetry operations such that for each $\mathcal{A}_n^{\text{local}}$ we arrive at an operator $\mathcal{A}_n$ that preserves all the relevant symmetries. If we start from all creation and annihilation operators for the orbitals in the cluster we can form all possible products of any number of them. Many of them don't preserve the number of particles and are irrelevant for our scheme. After we apply the symmetry projection some are linear combinations of others and can be dropped. Some linear dependence may not be obvious and may in fact also depend on the Hilbert space of the wavefunctions. This is the case for the particle number operator which is proportional to the unity operator. It occurs naturally if we start from an operator measuring the number of particles at one site of the cluster and we include the translations in the symmetries. An important point about these cluster operators is that there are a finite and in fact quite manageable number of them and that they can be constructed automatically. Of course the exact number of operators depends on the choice of the cluster and of the symmetries. For our method it is also important that this number is independent of the lattice size. The complexity of an operator, i.e., the number of individual terms in the sum, depends on the symmetries and grows linearly with the size of the lattice. This is in stark contrast to the Lanczos iterations where the complexity increases exponentially with the number of iterations. Replacing the Lanczos iterations with the cluster operators makes it possible to improve precisely the local correlations that we have otherwise no control over in the wavefunction.

The new scheme opens another possibility which goes beyond the mere minimization of the energy. Our wavefunction is now of the form

$$|\Psi_{\text{var}}(\mu, \lambda)\rangle = \mathcal{A}(\mu) |\Psi(\lambda)\rangle. \tag{9}$$

With the variational parameter(s) λ we control the long range behavior of the wavefunction whereas the parameter(s) μ are responsible for the short range part. We can now choose a
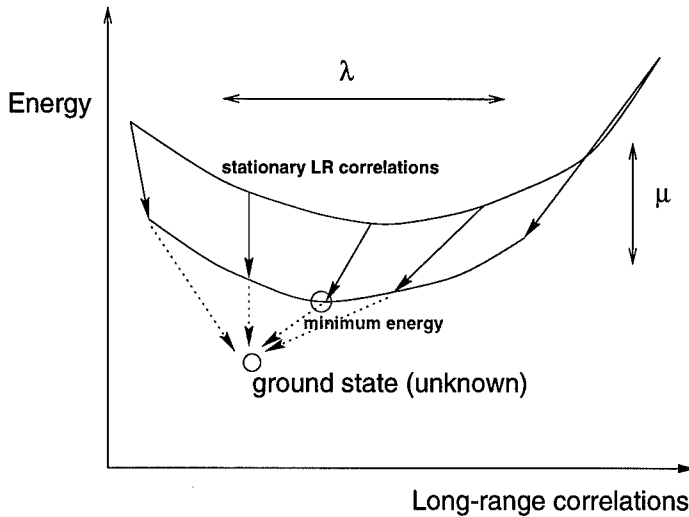
Figure 1: Schematic view of the convergence using cluster operators. Shown are the energy and a relevant measure for the long range correlations of the variational wavefunctions $\mathcal{A}(\mu)\,|\Psi(\lambda)\rangle$. The parameter λ controls the long range behavior and affects mainly the horizontal position while the parameter μ of the cluster operators accounts for some of the remaining correlation energy and moves the points down. The dotted lines suggest the expected convergence to the ground state if more complex cluster operators could be included in the calculations.

fixed λ and minimize the energy only by varying μ. Ideally this should leaves the long range properties invariant. We can now measure the corresponding long range correlation before and after the application of the operator $\mathcal{A}(\mu)$ and see whether this is really the case. If it should turn out that the short range operators are still changing the long range properties we can argue that the starting wavefunction $|\Psi(\lambda)\rangle$ did in fact not have the proper long range form. We can repeat this calculation with a different λ. By scanning through a number of values for λ we can find a wavefunction $|\Psi(\lambda)\rangle$ where the local operators don't affect the long range part when they optimize the energy. This wavefunction provides a good approximation for the long range properties of the ground state and only the local correlations need to be fixed. Figure 1 shows a schematic picture of this analysis. We measure both the energy and some property describing the long range correlation. For each λ there are two points, one corresponding to the bare $|\Psi(\lambda)\rangle$, and one for the optimized $\mathcal{A}(\mu_{\mathrm{optimal}}(\lambda))\,|\Psi(\lambda)\rangle$. The best energy estimate is of course still the lowest overall value and it is always an upper variational bound for the ground state energy. However, with the argument presented above the best estimate for the long range properties is the one that corresponds to the pair of points which shows no change in long range correlations. These points are in general not the same than those that lead to the lowest energy. The dotted lines show the expected convergence to the (generally unknown) ground state if we could extend the calculations to ever larger clusters for the cluster operators.

While the method just described is not mathematically rigorous, it provides a powerful
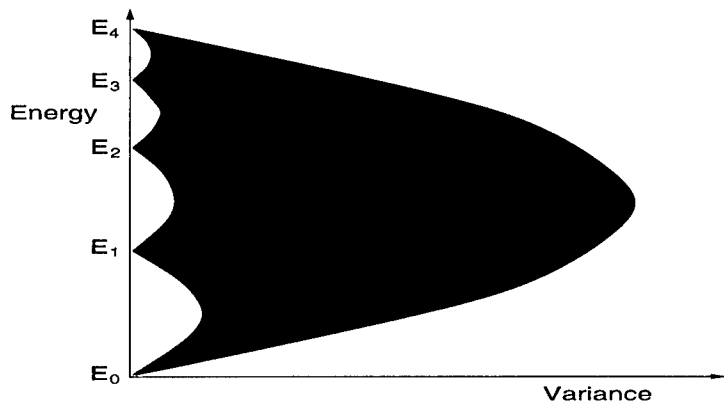
Figure 2: Region of possible values of the energy $\langle \mathcal{H} \rangle$ and variance $\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2$ for the variational wavefunctions. The convergence to the ground state will cause the variance to become smaller. Starting from a point close to an excited state will cause the variance to increase first when minimizing the energy. This provides an additional check for the quality of the variational wavefunction.

tool resting on physical intuition. Moreover there is still one more test which can be implemented easily. It involves the variance of the Hamiltonian. Suppose we know the spectral decomposistion of a wavefunction $|\Psi\rangle = \sum_{i=0}^{N} \alpha_i |\phi_i\rangle$ where $|\phi_i\rangle$ are the energy eigenstates corresponding to the energy E_i. Then we can calculate the energy and variance of that wavefunction

$$E = \langle \mathcal{H} \rangle = \sum_{i=0}^{N} \rho_i E_i, \tag{10}$$

$$\sigma_{\mathcal{H}}^2 = \langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 = \sum_{i=0}^{N} (\rho_i - \rho_i^2) E_i^2 - 2\sum_{i>j}^{N} \rho_i \rho_j E_i E_j. \tag{11}$$

If only two eigenstates are involved the points on an energy–variance plot lie on a parabola. In general there is a whole range of possible values and the corresponding region on the energy–variance plot is bounded by parabolas touching the energy axis at the eigenvalues as shown in Figure 2. It should be noted that the involved energies are only those which are compatible with the chosen symmetry properties and therefore the points on the energy axis may remain sparse even for bigger lattices. We see that the variance should converge to zero when approaching the ground state and in fact any energy eigenstate. If our starting wavefunction happens to be close to an excited state then the cluster operators are able to reduce the energy but they will increase the variance. Demanding that the variance is reduced at the same time that we minimize the energy provides us with a further test for the quality of the starting wavefunction.

RESULTS

The simplest model to describe strong correlations in a one-band tight-binding Hamiltonian is the Hubbard model. The Hubbard-Hamiltonian includes on-site Coulomb repulsion

with a coupling constant U and neglects interactions at a larger distance. In the strong coupling limit U becomes the dominant energy scale. The states can be characterized by the number of doubly occupied sites. If we are only interested in strong correlations and we want to describe electron densities of one electron per site or less we can restrict ourselves to the subspace with no double occupancy. This leads us to the t-J model as an effective Hamiltonian (*e.g.*, see [8])

$$\mathcal{H} = -t \sum_{<i,j>\sigma} (\tilde{c}_{i\sigma}^\dagger \tilde{c}_{j\sigma} + h.c.) + J \sum_{<i,j>} (\mathbf{S}_i \mathbf{S}_j - \frac{1}{4}n_i n_j) \tag{12}$$

with the usual convention of summing over nearest neighbor bonds $< i,j >$. The modified creation and annihilation operators prevent the formation of doubly occupied sites, *i.e.*, $\tilde{c}_{i\sigma} = c_{i\sigma}(1 - c_{i,-\sigma}^\dagger c_{i,-\sigma})$. They don't obey Fermi commutation relations and there is no weak coupling limit of the t-J model. While this makes the model more complex for most analytical treatments it turns out to be simpler for most numerical work. In recent years it has been the basis a lot of numerical work [9].

For large values of J/t the 2-dimensional t-J model shows phase separation between an antiferromagnetic Heisenberg ordered phase and one with a low electron density. In this way the system tries to minimize the number of broken bonds and we obtain an effective interaction between the holes. Exact diagonalization of small cluster suggests that this leads to a ground state with d-wave superconducting order for intermediate values of J/t [10]. Hence we choose for the starting point of our variational method a BCS-type wavefunction composed of Cooper pairs with d-wave symmetry [5]:

$$|\Psi(\lambda)\rangle = \mathcal{P}_G \mathcal{P}_N \prod_{\mathbf{k}\in\mathrm{BZ}} \left(u_\mathbf{k} + v_\mathbf{k} c_{\mathbf{k},\uparrow}^\dagger c_{-\mathbf{k},\downarrow}^\dagger \right) |vacuum\rangle \, ,$$

$$\begin{aligned} v_\mathbf{k}/u_\mathbf{k} &= \Delta_\mathbf{k}/\left(\epsilon_\mathbf{k} + \sqrt{\epsilon_\mathbf{k}^2 + \Delta_\mathbf{k}^2}\right), \\ \Delta_\mathbf{k} &= \lambda(\cos(k_x) - \cos(k_y)), \\ \epsilon_\mathbf{k} &= -2t(\cos(k_x) + \cos(k_y)) - \epsilon_\mathrm{F} \end{aligned} \tag{13}$$

The Gutzwiller projector $\mathcal{P}_G$ projects out all configurations with double occupancy and $\mathcal{P}_N$ is the N-particle projector. The variational parameter is λ. For $\lambda \to 0$ the variational wavefunction reduces to the Fermi sea (up to the Gutzwiller projection) and any finite value leads to superconducting long range order. The chemical potential ϵ_F does not need to be adjusted very accurately because $\mathcal{P}_N$ makes sure that we work with a fixed particle number. The cluster operators are constructed using a two by two cluster and enforcing the translational and rotational symmetries of the lattice as well as the spin rotation symmetry. The calculations are done on a 50 sites lattice with periodic boundary conditions along the (7,1) and (-1,7) directions. This lattice has the advantage that no reciprocal lattice vector $\mathbf{k}$ lies on the diagonal where $\Delta_\mathbf{k}$ vanishes. We choose to investigate the fillings of 0, 8, 12, 16, 20, and 24 holes relative to the half filled case of one electron per site. In these cases the variance of the Hamiltonian is consistently reduced when minimizing the energy thus suggesting that we are converging to the ground state. A test calculation for four holes has shown that the variance increases when minimizing the energy. In this case the variational wavefunction has too much overlap with low-lying excited states and is not appropriate for our method. For this reason we include only results for the hole densities mentioned above in the following presentation. Some preliminary results can be found in [11].

In order to measure long range correlations we use the pair-pair correlation function

$$C(R) = \frac{1}{N} \sum_i \langle \Delta^\dagger(i)\Delta(i+R)\rangle \tag{14}$$

with the Cooper pair annihilation operator

$$\Delta(i) = \frac{1}{2} \sum_\sigma c_{i,-\sigma}(c_{i+\hat{x},\sigma} + c_{i-\hat{x},\sigma} - c_{i+\hat{y},\sigma} - c_{i-\hat{y},\sigma}). \tag{15}$$

Our measure for long range correlations is $C_\infty = C(R_{max})$, $i.e.$, the pair-pair correlation at the largest distance. Figure 3 shows the energy versus the long range correlations for a hole density of 0.16 corresponding to 8 holes in 50 sites. Shown are the results for three different values of J/t. We see that for $J/t = 0.2$ there is no long range order and C_∞ is suppressed. Superconducting long range order only appears at higher values of J/t. We can collect the values C_∞ where the cluster operator don't change the long range order and plot them versus J/t as shown in figure 4. Here we can clearly see the onset of superconducting order at a finite value of $J_s = (0.39 \pm 0.03)t$ for a hole doping of $\delta = 0.16$. The same analysis is repeated for the other dopings. A test calculation for a 16 sites lattice shows excellent agreement with exact calculations.

Before we combine our results into a phase diagram let us have a look at phase separation. For very large values of J/t the system will try to optimize the magnetic energy by forming an antiferromagnetically ordered phase without holes thus leaving the holes in a dilute phase[12]. A standard Maxwell construction shows whether a combination of two phases with different densities has a lower energy than a homogeneous phase at the same total density. Our variational calculation gives a good estimate for $E(\delta)$ for several hole densities δ and allows us to construct an interpolation. If this function bends downward we can find a straight line connecting the phases at two densities. This corresponds to the energy of the phase separated state and we can find the region where phase separation has a lower energy. In this region our homogeneous variational wavefunction does not describe the ground state and the results for the superconducting order do not apply. The region of superconducting order lies between the onset of the superconducting order and the formation of the phase separated state. Figure 5 shows the resulting phase diagram of hole density δ versus coupling constant J/t. There is still some controversy about where the line of phase separation lies for low hole densities. Our results agree well with some recent calculations obtained by variational methods[13] but presently it remains an open question whether phase separation persists down to $J/t = 0$ as suggested by some Green's function Monte Carlo calculations[14].

CONCLUSION

We have shown that we can extend the traditional variational scheme to reliably measure quantities other than the ground state energy. Specifically it is possible to obtain results for long range correlations which relate directly to some order parameter. The key feature of the new scheme is a systematic separation of variational parameters controlling the long range versus the short range behavior. While the energy is still obtained as an upper variational bound according to the Rayleigh-Ritz principle, the long range correlation is measured by a new stationarity condition. This stationarity condition demands that the long range correlations be unaffected when optimizing for the short range part
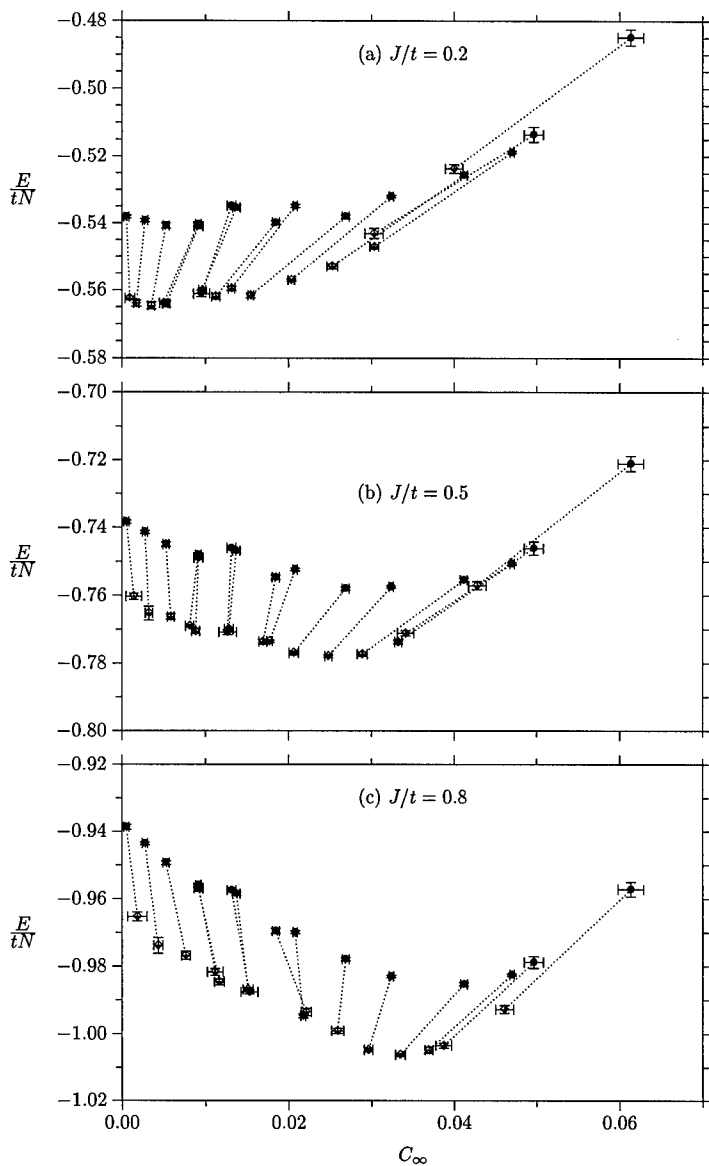
Figure 3: Energy versus long range correlations before and after minimization with the cluster operators for a hole density of $\delta = 0.16$. The plots show results for three representative coupling constants $J/t = 0.2, 0.5, 0.8$. The value for which the long range correlations C_∞ remains unchanged before and after minimization starts close to zero and increases monotonically with J/t.
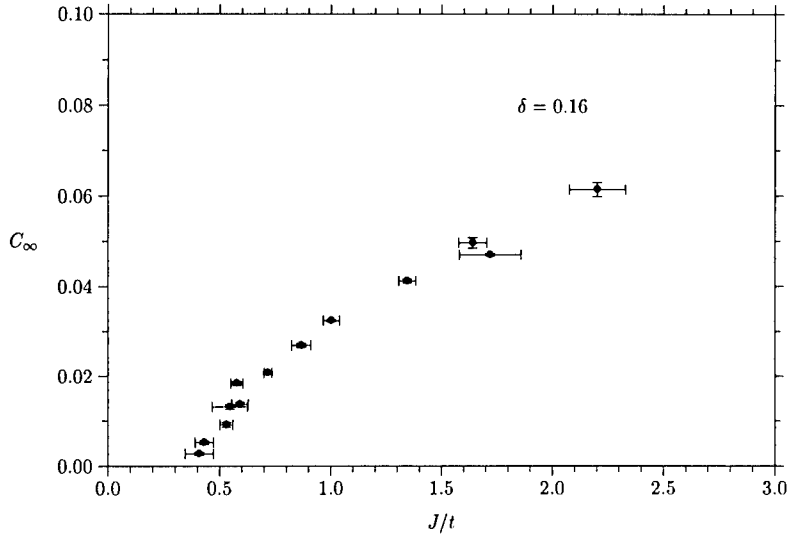
Figure 4: Long range d-wave correlation C_∞ as a function of coupling constant J/t for a hole density of $\delta = 0.16$. The points denote the values measured by the stationarity condition of our extended variational scheme.
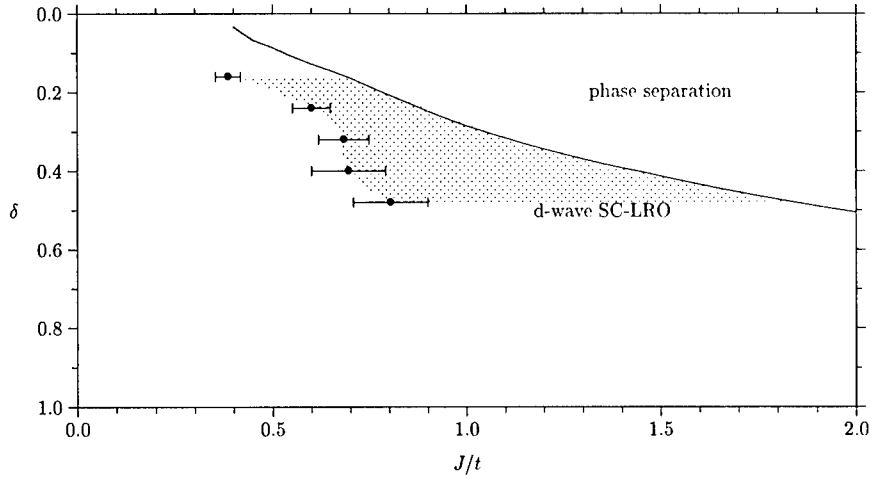
Figure 5: Phase diagram of the t-J model as a function of hole density δ versus coupling constant J/t. The solid line indicates the transition to a phase separated state. The phase separation line is obtained by a Maxwell construction using the homogeneous variational states. Below the phase separation the measured superconducting long range order indicates a region of a superconducting ground state. The points denote the onset of superconductivity for the densities under investigation.

of the wavefunction. Some other modified variational approaches like the power-Lanczos method[15] or the use of Gutzwiller-Jastrow type modfications[16] sometimes allow one to obtain better energy estimates but they remain weak in addressing the long range order. Our new method allows us to measure the long range order and thus to investigate the region of superconductivity of the ground state in the t-J model. We find that there is an extended region of hole dopings and coupling constants in the phase diagram of the t-J model where a superconducting ground state is most likely.

ACKNOWLEDGEMENT

REFERENCES

1. J. Hubbard, Proc. Roy. Soc. (London), Ser. A **276**, 238 (1963). J. Hubbard, Proc. Roy. Soc. (London), Ser. A **281**, 401 (1964).

2. E. H. Lieb and F. Y. Wu, Phys. Lett. **20**, 1445 (1968).

3. J. Sólyom, Adv. Phys. **28**, 201 (1979).

4. H. Yokoyama and H. Shiba, J. Phys. Soc. Japan **56**, 3570 (1987). C. Gros, Ann. Phys. (N.Y.) **189**, 53 (1989) and references therein.

5. Q. P. Li, B. E. C. Koltenbah, and R. Joynt, Phys. Rev. B **48**, 437 (1993), and references therein.

6. D. Ceperley, G. V. Chester, and M. H. Kalos, Phys. Rev. B **16**, 3081 (1977).

7. E. S. Heeb and T. M. Rice, Z. Phys. B **90**, 73 (1993).

8. F. C. Zhang and T. M. Rice, Phys. Rev. B **37**, 3759 (1988).

9. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).

10. E. Dagotto and J. Riera, Phys. Rev. Lett. **70**, 682 (1993).

11. E. S. Heeb and T. M. Rice, Europhys. Lett. **27**, 673 (1994).

12. V. J. Emery, S. A. Kivelson, and H. Q. Lin, Phys. Rev. Lett. **64**, 475 (1990). W. O. Putikka, M. U. Luchini, and T. M. Rice, Phys. Rev. Lett. **68**, 538 (1992).

13. C. T. Shih, Y. C. Chen, and T. K. Lee, preprint (cond-mat/9705156). M. Kohno, Phys. Rev. B **55**, 1435 (1997).

14. C. S. Hellberg and E. Manousakis, Phys. Rev. Lett **78**, 4609 (1997).

15. Y. C. Chen and T. K. Lee, Phys. Rev. B **51**, 6723 (1995).

16. H. Yokoyama and M. Ogata, preprint (cond-mat/9607050).

ON-SITE CORRELATION IN NARROW BAND MATERIALS

F. MANGHI*, V. BELLINI*, M. RONTANI*, C. ARCANGELI**
*Istituto Nazionale per la Fisica della Materia and Dipartimento di Fisica,Università di Modena, Via Campi 213/a, I-41100 Modena, Italy
** Max-Planck-Institut für Festkörperforschung, Heisenbergstr. 1, D-70569 Stuttgart, Germany

ABSTRACT

We present the results of a recently developed approach where the interplay between the itinerant and localized character of electrons in narrow band materials is described by adding on-site correlation effects to a realistic band calculation: the single particle band states are treated as mean field solutions of a multi-orbital Hubbard Hamiltonian and the many-body term associated with localized e-e interaction is described in a configuration-interaction scheme. Quasi-particle states of nickel and CuGeO3 have been calculated and compared with spectroscopical results.

INTRODUCTION

The Hubbard model, dominated by the competition between inter-site hopping and on-site electron-electron repulsion, is believed to describe the physics of narrow band materials such as transition metals, transition metal oxides, cuprates, etc.. In these systems the itinerant character of valence electrons shown by the k-dispersion observed in photoemission coexists with strong local correlations responsible for other spectroscopical features such as satellites, band-narrowing, and opening, in some cases, of a Mott-Hubbard gap. In spite of the enormous amount of work which has been done on cuprates since the discovery of high T_c superconductors, a unified theoretical description of the whole valence spectrum, from the high binding energy region characterized by satellites, up to the valence band top, including both unperturbed single particle like and strongly correlated Cu derived structures, is still missing; this is due to the difficulty to combine an accurate treatment of many body terms with a realistic description of the band structure.

Most of the work on this subject has been based on a drastic simplification of either the band structure or the e-e interaction [1]; describing the solid as a finite cluster the e-e intraction can be treated accurately, for instance by exact diagonalization techniques but in the search for the simplest model containing the relevant physics of superconductors one may miss some important effects and the possibility of a quantitative comparison with spectroscopical results. Photoemission data of highly correlated materials have been also interpreted using cluster configuration-interaction models which assume a strong wave function localization and adopt a rather simplified description of the band structure, with a considerable number of adjusting parameters [2]. Other approaches have been based on the density functional approximation (self-interaction corrected [3] and LDA+U [4] functionals) which fully include the itinerant character of electron states but describe the electron-electron interaction as a mean field effective single particle potential.

A theoretical approach is then needed which includes both the hybridization between Cu and the ligands (or between sp and d states in the case of transition metals) accounted for by first principle band theory, and a treatment of e-e interaction which must be non-

perturbative, to deal with systems which are in the high correlation regime, and beyond mean field, to include finite life-time excitations.

The 3 Body Scattering (3BS) method [5, 6] can be seen as an extension to the solid state of the configuration-interaction scheme used for finite systems: the Hubbard Hamiltonian is projected on a set of states obtained by adding a finite number of e-h pairs to the ground state of the single particle Hamiltonian and this expansion is truncated to include one e-h pair. The effect of electron correlation on one-electron removal energies from a partially filled band is then described as hole-hole and hole-electron interaction. The 3BS theory corresponds to the solution of a 3-body scattering problem involving two holes and one electron. Self-energy corrections, spectral functions and quasi-particle band structure can be calculated for systems in different correlation regimes, getting a complete picture of the whole valence spectrum, including both long-lived coherent quasiparticle structures and incoherent short-lived ones.

THEORY

Both Hubbard and single particle band Hamiltonians are approximate descriptions of the interacting system which differ for the treatment of the on-site e-e repulsion. They can be formally obtained applying mean field approximation to the exact many body Hamiltonian, either to all the many body terms, or selectively to the multi-center integrals getting an effective single-particle problem ($\hat{H}^{BND}$), or the generalized Hubbard model ($\hat{H}^{H}$), respectively. In a Bloch basis set the two approximate Hamiltonians are

$$
\hat{H}^{H} = \sum_{kn\sigma} \epsilon_{kn}^{H} \hat{a}_{k\sigma}^{n\dagger} \hat{a}_{k\sigma}^{n} + \sum_{\substack{\alpha\beta \; nn' \\ kk'p \; mm'}} \frac{U_{\alpha\beta}}{N} C_{\alpha\uparrow}^{n}(k)^{*} C_{\alpha\uparrow}^{n'}(k+p)
$$

$$
\times \quad C_{\beta\downarrow}^{m}(k')^{*} C_{\beta\downarrow}^{m'}(k'-p) \hat{a}_{k\uparrow}^{n\dagger} \hat{a}_{k+p\uparrow}^{n'} \hat{a}_{k'\downarrow}^{m\dagger} \hat{a}_{k'-p\downarrow}^{m'},
$$

$$
\hat{H}^{BND} = \sum_{kn\sigma} \epsilon_{kn\sigma}^{BND} \hat{a}_{k\sigma}^{n\dagger} \hat{a}_{k\sigma}^{n}; \tag{1}
$$

here $\hat{a}_{k\sigma}^{n}, \hat{a}_{k\sigma}^{n\dagger}$ are destruction/creation operators of electrons with wave vector k, spin σ, band index n, $C_{\alpha\sigma}^{n}(k)$ are the expansion coefficients of Bloch states in terms of localized orbitals and $U_{\alpha\beta}$ is the on-site Coulomb repulsion among orbitals α and β. The relationship between single particle eigenvalues $\epsilon_{kn\sigma}^{BND}$ and $\epsilon_{kn\sigma}^{H}$ appearing in the two Hamiltonians is

$$
\epsilon_{kn\sigma}^{BND} = \epsilon_{kn\sigma}^{H} + Q_{k\sigma}^{n}, \tag{2}
$$

$$
Q_{k\sigma}^{n} = \sum_{\alpha\beta} |C_{\alpha\sigma}^{n}(k)|^{2} \left[U_{\alpha\beta} \frac{1}{N} \sum_{k'n'}^{occ} |C_{\beta-\sigma}^{n'}(k')|^{2} \right]. \tag{3}
$$

Since we want to augment band theory with the inclusion of on-site correlation, equations (2,3) are essential to define the correct relationship between band and Hubbard Hamiltonian, and to avoid double counting of e-e interaction.

We are interested in the hole spectral function

$$
D^{-}(\omega) = \sum_{kn} D_{kn}^{-}(\omega),
$$

$$
D_{kn}^{-}(\omega) = \frac{1}{\pi} \mathrm{Im} \frac{1}{\omega - \epsilon_{k}^{n} - \Sigma_{nk}^{-}(\omega)},
$$

which describes the response of the system to one-electron removal and is the quantity directly related to the photoemission results.

In order to calculate the hole self-energy $\Sigma_{n\mathbf{k}}^-$ we proceed as in ref. [7], adopting a configuration-interaction scheme which consists in projecting the Hubbard Hamiltonian on a set of states obtained by adding a finite number of e-h pairs to the Fermi sea, i.e. to the ground state of the single particle Hamiltonian. The theory has been formally developed to include configurations with up to two e-h pairs [6] but in the limit of an almost filled band, as it is the case for the systems of interest here, the configurations can be usefully reduced to those involving one e-h pair only; in this way the problem remains an extension of previous approaches based on a simplified description of the scattering channels (the so called T-matrix approaches [8]) but it is more tractable.

The interactions between the 3-body configurations (one hole plus one e-h pair) are represented by a set of scattering T-matrices, describing h-h scattering

$$T_{\alpha\beta}^{h-h}(\omega) = \frac{U_{\alpha\beta}}{1 + U_{\alpha\beta}g_{\alpha\beta}^{h-h}(\omega)}, \tag{4}$$

with

$$g_{\alpha\beta}^{h-h}(\omega) = \int_{-\infty}^{E_f} d\,\epsilon' \int_{-\infty}^{E_f} d\,\epsilon \frac{n_\alpha(\epsilon)n_\beta(\epsilon')}{\omega - \epsilon' - \epsilon - i\delta}, \tag{5}$$

and e-h scattering

$$T_{\alpha\beta}^{h-e} = \frac{-U_{\alpha\beta}}{1 - U_{\alpha\beta}g_{\alpha\beta}^{h-e}(\omega)}, \tag{6}$$

with

$$g_{\alpha\beta}^{h-e}(\omega) = \int_{-\infty}^{E_f} d\,\epsilon' \int_{E_f}^{\infty} d\,\epsilon \frac{n_\alpha(\epsilon)n_\beta(\epsilon')}{\omega - \epsilon' + \epsilon - i\delta}; \tag{7}$$

$n_\alpha(\epsilon)$ is the orbital density of single-particle band states.

The Faddeev theory [6] is used to determine the total scattering matrix and the resolvent of the many body system. The hole self-energy is given by

$$\Sigma_{n\mathbf{k}}^-(\omega) = \sum_\beta |C_\beta^n(\mathbf{k})|^2 \left[\sum_\alpha U_{\alpha\beta} N_\alpha^h - \Sigma_\beta^-(\omega) \right], \tag{8}$$

where N_α^h is the percentage of empty states in the orbital α and

$$\Sigma_\beta^-(\omega) = \sum_\alpha \int_{E_f}^{\infty} n_\alpha(\epsilon)T_{\alpha\beta}^{h-h}(\omega - \epsilon)\left[1 + U_{\alpha\beta}A_{\alpha\beta}(\omega - \epsilon)\right] d\,\epsilon;$$

$A_{\alpha\beta}$ is the quantity related to $T_{\alpha\beta}^{h-e}$ and is determined solving numerically the integral equation described in ref. [6, 7].

We have applied this method to a transition metal (nickel) and a cuprate ($CuGeO_3$) assuming the localized e-e repulsion among d-states as the dominant contribution, i.e.

$$U_{\alpha\beta} = \begin{cases} U_{dd} & \text{for } \alpha, \beta = d \text{ orbitals} \\ 0 & \text{elsewhere;} \end{cases}$$

$$U_{\alpha\beta} - J_{\alpha\beta} \simeq 0.$$

To apply this method to $CuGeO_3$ we have used the eigenstates/eigenvalues of ref. [9] and assumed $U_{dd} = 8$ eV. In fig. 1 we plot k-resolved spectral functions along high symmetry directions and compare them with single particle band states.
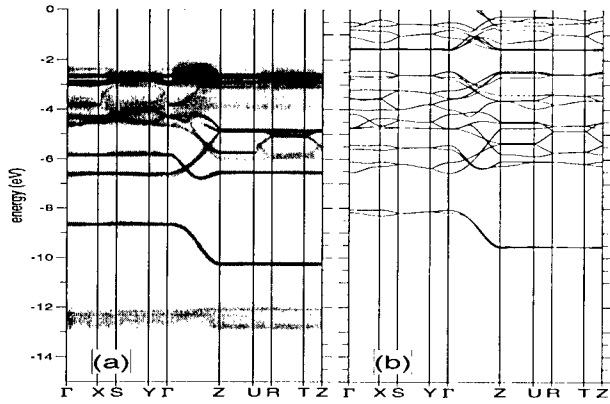
Figure 1: (a) Grey scale representation of the k-resolved hole spectral functions of paramagnetic $CuGeO_3$ with black as maximum intensity. (b) Single-particle band structure of ref. [9].

The peaks in the spectral function can be classified either as quasiparticle excitations or as satellites, according to their intrinsic line-width arising from different values of the imaginary part of self-energy: quasi-particle excitations (sharp lines in fig. 1) correspond to a small imaginary part, and give rise to the coherent part of the spectral function. Satellites (smeared lines) occur where the imaginary part of self-energy is large and correspond to short-lived excitations with a large intrinsic line-width; we refer to them as the incoherent part of the spectral function. The effect of electron correlation on single particle states is dramatic: some bands are shifted to higher binding energies, spectral weight is removed from the upper part of the spectrum, and many new states (satellites) appear; only states around -8, -9 eV and -5, -6 eV are practically unaffected being mainly Ge and O derived.

$CuGeO_3$ is an insulator but it is predicted to be a metal by single particle band calculation; the inclusion of electron correlation is able to reproduce its insulating behaviour. The same was proven to be true also in the case of NiO, where a 3BS description of Hubbard correlation was able to reproduce both the complex satellite structure and the measured value of the insulating gap [7].

The ability of 3BS approach to open up Hubbard gaps, i.e. to reproduce an insulating behaviour in a system which is metallic according to band theory, is related to its non-perturbative character; in fact it has been shown [5, 7] that for U much larger than the band width W ($\frac{U}{W} \to \infty$) 3BS reproduces exactly the so called "atomic limit" solution of

Hubbard Hamiltonian, where hole and electron states are separated by a Mott-Hubbard gap equal to U.

In the case of nickel the on-site e-e repulsion is more effectively screened and the estimated value of U_{dd} is $\simeq 2$ eV [10]. Fig. 2, reporting the comparison between quasi particle states and single particle ones, shows that e-e correlation effects are still sizable and they are actually essential in order to reproduce the observed spectroscopical features, i.e. satellite structure at 6 eV binding energy, correct band width (overestimated in LDA), exchange splitting [11], and energy dispersion [12].
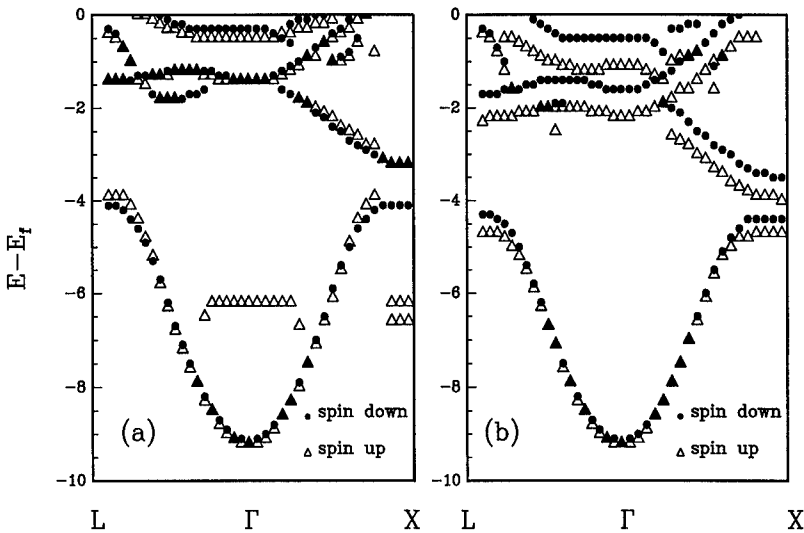

Figure 2: (a) Energy position of the peaks of nickel in the spectral function at k-points along high symmetry lines of the Brillouin Zone. (b) Single particle band structure.

SUMMARY AND OUTLOOK

We have described a method to include on-site interactions in the description of hole and electrons states: ab-initio single particle band states are used as input mean field eigenstates for the calculation of self-energy corrections according to a 3-body scattering (3BS) solution of a multi-orbital Hubbard Hamiltonian. When applied to valence states of ferromagnetic nickel it allows to get a quasiparticle band structure which compares much more favorably with the experimental observation than conventional mean field LDA, reproducing the observed band width, the energy dispersion, the satellite structure and the exchange splitting. Since the method does not rely on a perturbation expansion it has a wide range of application, including any correlation regime. In the case of a highly

correlated system such as $CuGeO_3$, 3BS is able to reproduce both the insulating behaviour and a correct overall picture of photoemission experiments.

Our present choice of empirically determining the parameter U of the Hubbard Hamiltonian - which has been fixed to reproduce the satellite binding energy - ensures that we obtain a good agreement with experiments; however, previous T-matrix methods [13, 14] have not been able to reproduce at the same time the satellite energy position of nickel and the valence band width which turned out to be systematically overestimated for values of the Coulomb integral fixed to reproduce the satellite binding energy. The possibility of reproducing both the satellite structures and other important spectroscopical features (band width and energy dispersion in nickel, the insulating gap in $CuGeO_3$) can be seen as a non trivial result and a success of the method itself. The problem of extracting Hubbard U from an ab-initio calculation, either in the so called Constrained-Density Functional scheme [15] or as screened Coulomb interaction [16], is an important issue which goes in the direction of a full match between model Hamiltonians and realistic systems and that we are presently considering as an implementation of our approach.

REFERENCES

1. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).

2. H. Eskes, L. H. Tjeng, and G. Sawatzky, Phys. Rev. B **41**, 288 (1990).

3. Axel Svane, Phys. Rev. Lett. **68**, 1900 (1992).

4. V. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).

5. J. Igarashi, J. Phys. Soc. Japan **52**, 2827 (1983); *ibidem* **54**, 260 (1985).

6. C. Calandra and F. Manghi, Phys. Rev. B **50**, 2061 (1994).

7. F. Manghi, C. Calandra, and S. Ossicini, Phys. Rev. Lett. **73**, 3129 (1994).

8. J. Kanamori, Progr. Theoret. Phys. **30**, 275 (1963)

9. L. F. Mattheiss, Phys. Rev. B **49**, 14050 (1994).

10. M. Springer and F. Aryasetiawan, to be published.

11. T. J. Kreutz, P. Aebi, and J. Osterwalder, Solid State Commun. **96**, 339 (1995).

12. F. Manghi, V. Bellini, and C. Arcangeli, Phys. Rev. B **56**, 7149 (1997).

13. A. Liebsch, Phys. Rev. Lett. **43**, 1431 (1979).

14. C. Calandra and F. Manghi, Phys. Rev. B **45**, 5819 (1992).

15. P. H. Dederichs, S. Bluegel, R. Zeller, and H. Akai, Phys. Rev. Lett. **53**, 2512 (1984).

16. M. Rontani, F. Rossi, F. Manghi, and E. Molinari, submitted.

THREE-BODY CORRELATION IN THE DILUTED GENERALIZED HUBBARD MODEL

O. NAVARRO
Instituto de Investigaciones en Materiales, U.N.A.M.,
Apartado Postal 70-360, 04510, México D.F.,
MEXICO.

M. AVIGNON
Laboratoire d'Etudes des Propiétés Electroniques des Solides, C.N.R.S.,
Boîte Postale 166, 38042 Grenoble Cedex 9,
FRANCE.

ABSTRACT

A real-space method has been used to solve the generalized Hubbard Hamiltonian for a system with few electrons. The method is based on mapping the correlated many-body problem onto an equivalent tight-binding one in a higher dimensional space. For a linear chain, we have obtained an exact solution of the problem of three non-parallel electrons. The three-body correlation are studied by examining the binding energy in the ground state, for different values of the hopping parameters and of the on-site (U) and nearest-neighbor (V) interactions.

INTRODUCTION

Since the discovery of high-T_c superconductivity [1] in $Cu - O$ compounds, the studies on the physics of correlated electrons in low dimensional systems have become extremely important. Although the main interest is on the physics of two-dimensional highly correlated electron systems, the one-dimensional models related to high temperature superconductivity are very popular due to the conjecture [2] that properties of the 1D and 2D variants of certain models have common aspects. Within the models for correlated electron systems, that attempt to capture the essential physics of high-temperature superconductors and parent compounds, the simple Hubbard model [3], is the crudest approximation to include electronic interaction between band-electrons, by retaining only the on-site interaction U. This model also assigns the same hopping rate t to three different hopping processes regardless of the occupation of the two sites involved.

Besides the on-site interaction, other contributions of the electron-electron interaction are required [4], such as the nearest-neighbor interactions and the bond-charge interaction term. The Hamiltonian which includes these interactions is often called the generalized Hubbard Hamiltonian and has been studied previously by several authors [5-8]. This Hamiltonian can be written as

$$H = \sum_{<i,j>,\sigma} t_{i,j}^{\sigma}(c_{i,\sigma}^{+}c_{j,\sigma} + h.c.) + U\sum_{i} n_{i,\uparrow}n_{i,\downarrow} + \frac{V}{2} \sum_{<i,j>} n_i n_j, \tag{1}$$

185

where $< i, j >$ denotes nearest-neighbor sites, $c_{i,\sigma}^{+}$ $(c_{i,\sigma})$ is the creation (annihilation) operator with spin $\sigma = \downarrow$ or $\uparrow$ at site i, and $n_i = n_{i,\uparrow} + n_{i,\downarrow}$ where $n_{i,\sigma} = c_{i,\sigma}^{+} c_{i,\sigma}$. It is worth mention that in principle, the parameters U and V are positives because they are direct Coulomb integrals. However, U and V could be negative if attractive indirect interaction through phonons or other bosonic excitations are included and they are stronger than the direct Coulomb repulsion. In Eq. (1), the occupation-dependent hopping amplitude, $t_{i,j}^{\sigma}$, is given by

$$t_{i,j}^{\sigma} = t_{AA}(1 - n_{i,-\sigma})(1 - n_{j,-\sigma}) + t_{BB} n_{i,-\sigma} n_{j,-\sigma} +$$

$$t_{AB}[n_{j,-\sigma}(1 - n_{i,-\sigma}) + n_{i,-\sigma}(1 - n_{j,-\sigma})].$$

The three parameters t_{AA}, t_{BB}, and t_{AB} are the hopping amplitudes from a singly occupied to an empty site, from a doubly occupied to a singly site and from a doubly occupied to an empty site, respectively. The special case $t_{AA} = t_{BB} = t_{AB} = t$ corresponds to the t-U-V extended Hubbard model, which has been studied intensively by analytical and numerical methods [3]. When $t_{AA} + t_{BB} - 2t_{AB} = 0$ and $t_{AA} \neq t_{BB}$, the generalized model given by Eq. (1) reduces to the Hirsch and Marsiglio model of hole superconductivity [4]. In this paper, parings between electrons in a one-dimensional lattice using the generalized Hubbard Hamiltonian are analyzed. The analysis has been done following a mapping method previously reported [9].

RESULTS AND DISCUSSION

In this section, we study the correlation of three electrons by using the mapping method explained in Ref. 9. In order to present a brief explanation of the mapping method, let us consider the case of two electrons with opposite spins in an N-site chain; the number of states is given by N^2. This states form a square lattice with $(3N - 2)$ "impurities", which can be described by a single-body tight-binding Hamiltonian. Of these impurities, N are localized on sites along the principal diagonal of the square lattice with a self-energy U and the others, $2(N - 1)$, are localized on the two next-diagonals with a self-energy V. A simple way to obtain the solution is taking advantage of the translational symmetry of the impurities and projecting the two-dimensional lattice of states onto a linear chain of effective states, similar to the procedure introduced by Falicov and Yndurain [10]. In general, this method will map the original many-body problem onto a tight-binding one with some ordered impurities in an nd-dimensional lattice, being n the number of electrons and d the dimensionality of the original system. In this hyper-space lattice, the on-site (U) and the nearest-neighbor (V) interactions from the original Hubbard Hamiltonian become the self-energies of the impurities.

For the case of three electrons, two with up-spin and one with down-spin in a linear chain, the network of states belongs to a three-dimensional state lattice. For example, let us consider these three electrons in a 4-site chain. The state configuration is:

$|1\rangle = |\pm +00\rangle, |2\rangle = |\pm 0 + 0\rangle, |3\rangle = |\pm 00+\rangle, |4\rangle = |-++0\rangle,$
$|5\rangle = |-+0+\rangle, |6\rangle = |-0++\rangle, |7\rangle = |+\pm 00\rangle, |8\rangle = |+-+0\rangle,$
$|9\rangle = |+-0+\rangle, |10\rangle = |0\pm +0\rangle, |11\rangle = |0\pm 0+\rangle, |12\rangle = |0-++\rangle,$
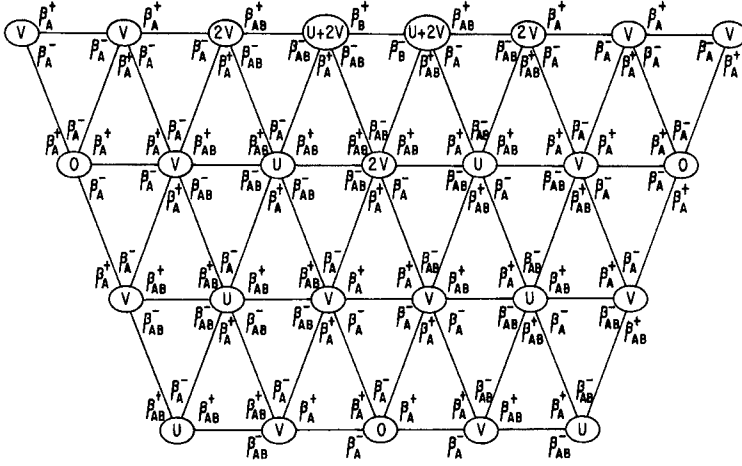
Figure 1: Lattice of effective states for three electrons with non-parallel spin in a linear chain. These effective states are represented by ellipses and the self-energy for each of them is indicated inside. There are six different effective-hopping parameters β_A^+, β_A^-, β_B^+, β_B^-, β_{AB}^+, and β_{AB}^-, with values given in the text.

$$|13\rangle = |++-0\rangle, |14\rangle = |+0\pm 0\rangle, |15\rangle = |+0-+\rangle, |16\rangle = |0+\pm 0\rangle,$$
$$|17\rangle = |0+-+\rangle, |18\rangle = |00\pm+\rangle, |19\rangle = |++0-\rangle, |20\rangle = |+0+-\rangle,$$
$$|21\rangle = |+00\pm\rangle, |22\rangle = |0++-\rangle, |23\rangle = |0+0\pm\rangle, |24\rangle = |00+\pm\rangle.$$

Spin up and down are denoted by $+$ and $-$, respectively, a doubly-occupied site by $\pm$, and a hole by 0. A site occupied by two electrons requires an energy U, and the amplitudes of the transition probability for nearest-neighbor states will be t_{AA}, t_{BB} and t_{AB}. In general, for an N-site chain the number of states is given by $(N-1)(3N-2)/2$. As we already mentioned, the geometric representation of these states belongs to a three-dimensional lattice, where taking advantage of the translational symmetry in this network of states, it can be projected onto a two-dimensional lattice of effective states.

In Fig. 1 we show the two-dimensional lattice of states for an original system with three electrons in a five-site chain, where the hopping parameters β_A^+, β_A^-, β_B^+, β_B^-, β_{AB}^+, and β_{AB}^- are given by

$$\beta_A^+ = t_A e^{iKa/\sqrt{3}}, \quad \beta_A^- = t_A e^{-iKa/\sqrt{3}},$$
$$\beta_B^+ = t_B e^{iKa/\sqrt{3}}, \quad \beta_B^- = t_B e^{-iKa/\sqrt{3}},$$
$$\beta_{AB}^+ = t_{AB} e^{iKa/\sqrt{3}}, \quad \beta_{AB}^- = t_{AB} e^{-iKa/\sqrt{3}}.$$

Here, K is the wave vector and a is the lattice parameter. The two-dimensional results must also be integrated with respect to K within the first Brillouin zone.
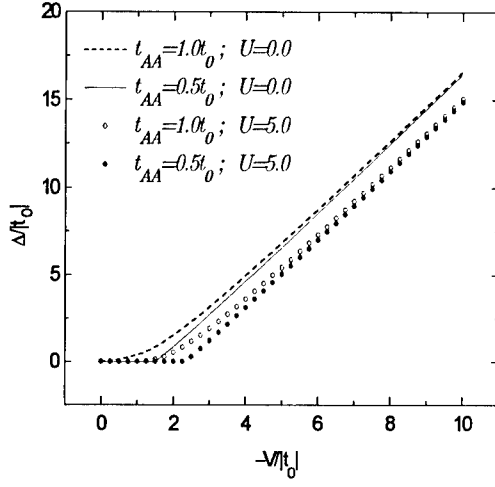
Figure 2: Binding energy (Δ) as a function of the nearest-neighbor atractive interaction ($-V$), for three non-parallel electrons in a linear chain. In this plot we made some variations of the t_{AA} hopping parameter for values of the on-site repulsion interaction term (U).
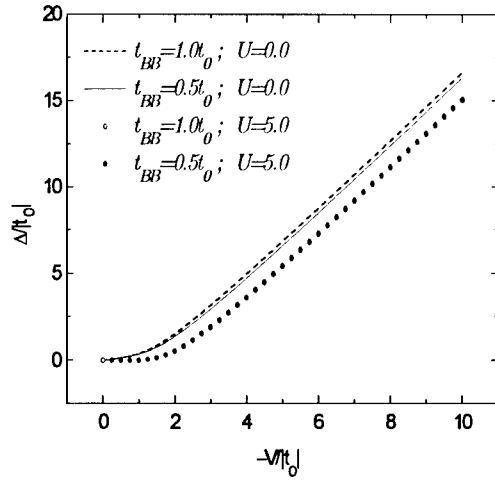

Figure 3: Same plot as in Fig. 2, but with variation on the t_{BB} hopping parameter.

The binding energy (Δ) as a function of the nearest-neighbor attractive interaction (V) is shown in Fig. 2 and Fig. 3, for the case of three non-parallel ($\uparrow\downarrow\uparrow$) electrons on a one-dimensional lattice. Here, the parameter t_0 has been taken to be equal to -1. The binding energy has been calculated from the energy difference between the lowest correlated state ($K = 0$), and the original lower band edge when there is not an electron-electron interaction. The final numerical diagonalization were carried out for a truncated two-dimensional lattice of 551 effective states. The matrix sizes for numerical diagonalizations were chosen as the minimum size so that the physical quantities, such as the binding energy, have not an important variation with the matrix size.

In Fig. 2, we show the binding energy for the case $t_{BB} = t_{AB} = t_0$, and for different values of the t_{AA} hopping parameters and the U repulsive interaction. We can observe that for small values of V, the binding energy is very sensitive to changes on the hopping probability t_{AA}, and that for very large values of this attractive interaction regime, Δ goes like V. It is worth mentioning that for the special case $t_{AA} = t_0$, our problem corresponds to a three-electron problem within the extended Hubbard Hamiltonian. The results for this case were previously calculated for $U = 0$ [11], and included here for comparison with our results.

In Fig. 3, we plot Δ for the case $t_{AA} = t_{AB} = t_0$, and for different values of t_{BB} and U parameters. For very strong attractive nearest-neighbor interactions the behavior is similar to the one given in Fig. 2, it means a linear behavior, while in the weak interaction regime the binding energy has and exponential behavior.

CONCLUSIONS

We have studied the electronic correlation in a generalized Hubbard model with hopping depending on the occupation, using a simple mapping method. This study has been carried out for a system with three-correlated electrons in a one-dimensional lattice, where an exact solution has been obtained for this problem. The electronic correlation is studied by examining the binding energy for differents values of the Hubbard parameters.

ACKNOWLEDGMENTS

REFERENCES

[1] J.G. Bednorz, K.A. Müller, *Z. Phys.* **B64**, 189 (1986).

[2] P.W. Anderson, *Science* **235** 1196 (1987); *Phys. Rev. Lett.* **64**, 1839 (1990).

[3] R. Micnas , J. Ranninger and S. Robaszkiewicz, Rev. Mod. Phys. **62** 113 (1990); *The Hubbard Model: Recent results*, edited by M. Rosetti, Series on Advances in Statistical Mechanics (World Scientific, Singapore, 1992) Vol. **7**.

[4] J. E. Hirsch and F. Marsiglio, Phys. Rev. B **41**, 2049 (1990).

[5] R. Strack and D. Vollhardt, Phys. Rev. Lett. **70**, 2637 (1993).

[6] A. A, Ovchinikov, Mod. Phys. Lett. B **7**, 21 (1993).

[7] L. Arrachea and A. A. Aligia, Phys. Rev. Lett. **73** 2240 (1994); Gagliano, A. A. Aligia, L. Arrache and M. Avignon, Phys. Rev B **51**, 14012 (1995).

[8] O. Navarro, Czech. J. Phys. **46,** S4 1867 (1996).

[9] O. Navarro and C. Wang, Solid State Commun. **83,** 473 (1992); L. A. Pérez, O. Navarro and C. Wang, Phys. Rev. B **53,** 15389 (1996); O. Navarro, *Introducción a la Superconductividad* (Aula Magna UAS, Vol. **11**, 1997).

[10] L. M. Falicov and F. Yndurain, J. of Phys. C: **8** 147 (1975).

[11] C. Wang, O. Navarro and R. Oviedo-Roa, *Mat. Res. Soc. Symp. Proc.* **291**, 279 (1993).

RELIABLE ESTIMATES OF QUASI-PARTICLE ENERGIES AND EXCI-TONIC EFFECTS IN CLUSTERS THROUGH DISCRETE-VARIATIONAL-METHOD TOTAL ENERGY CALCULATIONS

Giancarlo Cappellini and Francesco Casula

INFM and Dipartimento di Scienze Fisiche, Università di Cagliari, Italy

Friedhelm Bechstedt

FSU, Physikalisch-Astronomische Fakultät, Jena, Germany

ABSTRACT

Self-energy corrections to the DFT-LDA HOMO-LUMO gap for clusters have been calculated through an efficient method based on ΔSCF theory. Total energies for the clusters are obtained via discrete variational method (DVM). Good results in comparison with experiments and available theoretical data for the sodium tetramer Na_4 and C_{60} have been obtained. The localization properties of DVM basis-functions play a fundamental role in obtaining the above results with a strongly reduced computational effort in comparison with *ab-initio* schemes. For the two systems under study an estimate of the first excitonic transition energy is also given.

Introduction

Density-functional theory (DFT) in the local-density approximation (LDA) for the exchange and correlation energy functional has been successfully applied to determine the ground-state electronic properties of a large class of materials ranging from bulk systems, surfaces and etherostructures to atoms and clusters.[1] On the other hand if DFT-LDA is used to determine the quasiparticle (QP) spectra of many-electron systems in most cases results are obtained in disagreement with experiments.[1] For example, in the case of bulk semiconductors, assumption of the Kohn-Sham eigenvalues as electronic QP energies leads to a systematic underestimate of the electronic transition energies with respect to the experiment (*band-gap problem*).[2, 3]

In the present paper we apply the idea outlined by Gunnarson and Jones[1] in order to efficiently calculate the HOMO-LUMO (highest occupied molecular orbital-lowest unoccupied molecular orbital) gap and the QP spectra of clusters. It is based on the calculation of total DFT-LDA energies of charged and neutral systems. Since total energies only are needed, and not the full dielectric screening function of the system as in the case of the GW approach, [2, 3] the QP spectrum calculation is then feasible on any vax-class microcomputer. Total energies of the clusters in their ground-state are determined using a code founded on the discrete variational method (DVM).[4] This method has its basic idea in using a discrete sample of points for integrating in real space the hamiltonian and the overlap matrix elements. As benchmark for our method the fullerene molecule C_{60} and the sodium tetramer Na_4 have been choosen. They have respectively 240 and 4 *valence* electrons and represent some-ways the edge cases for electronic closed-shell clusters. The QP HOMO-LUMO gap for the former system results in good agreement with experiment, while for the second, the comparison with previous theoretical results and experiments turns out

HOMO-LUMO gap C$_{60}$	present work	previous results
DFT-LDA	1.79	1.87-2.18
DFT-LDA (fullerite)		1.04
HF		7.53
GW (fullerite)		2.15
I-A		4.70-5.20
E(1)	5.21	
E(2)	5.30	

Table I: HOMO-LUMO gap values in eV of neutral C$_{60}$ cluster as determined in this paper by standard DFT-LDA calculations and according to Eqs. [E(1)] and (2) [E(2)] are compared with previous theoretical and experimental results. The two results referring to solid fullerite are shown for the sake of completeness. The experimental range for the difference between vertical ionization potential (I) and electron affinity (A) has been reported. Details concerning previous results are given in Ref.[6]

to be less satisfactory. This difference can be ascribed to the different effects of the LDA in the two cases here considered. Calculations for C$_{60}$ and Na$_4$ have been performed at the equilibrium geometry of the neutral clusters. The fullerene C$_{60}$ case will be considered in the following in more detail due to the great relevance gained by this system in the last years.

I) The Buckminsterfullerene C$_{60}$

The discovery of C$_{60}$ molecule, of its condensation into an insulating solid and of superconductivity in the alkali-metal-doped solids have generated enormous interest in the fullerenes and fullerrites. We refer for an extended discussion on the electronic structure of fullerenes and their screening properties to Refs.[5, 6] and references therein. In Tab(I) we report data from the literature and after the present work concerning C$_{60}$. The previous theoretical and experimental results in the first five rows are referred and discussed in details in Ref.[6]. In the last two rows are the excitation energies determined after Eq.(1) -Eq.(3) in the present paper (see the following). As can be seen in this table the LDA HOMO-LUMO gap-values reported in the literature for the C$_{60}$ cluster, are more than 5.0 eV smaller than the Hartree-Fock (HF) one. Thus in C$_{60}$ the same *band-gap problem* occurring in bulk semiconductors and insulators takes place; while LDA underestimates on the other hand HF scheme overestimates the experimental band-gaps and electronic transition energies due to the different approximations used for the self-energy operator.[2, 3]

In Tab.(I) are also reported DFT-LDA and GW gaps calculated for the solid phase C$_{60}$ fullerite using a plane-wave basis set. The DFT-LDA one is smaller by about 1eV with respect to the corresponding gap of the cluster due to the higher confinement experienced by the extra electron in the cluster with respect to the solid phase. Analogously also the QP HOMO-LUMO gap for the isolated C$_{60}$ molecule is expected to be larger than that of the fullerite.

We calculate the QP HOMO-LUMO gap by using energy differences of charged and neutral C$_{60}$ clusters. The energy gap of a N electron system is rigorously defined within ΔSCF method as[1, 3]

$$E_{gap} = E_{N+1}^{TOT} + E_{N-1}^{TOT} - 2E_{N}^{TOT} \tag{1}$$

where E_{N}^{TOT} is the ground-state energy. To apply usefully DVM to C$_{60}$ a grid of 300 diophantine points has been chosen around each of the 60 equivalent atoms and a minimal

basis set has been chosen, made up of symmetrized combinations of 2s and 2p occupied carbon orbitals (1s frozen). Convergence with respect to integration-point number as well to additional unoccupied basis-orbitals has been checked here, and was also proved correct in previous calculations for similar materials.[4] Selfconsistency is achieved with respect to orbital occupation numbers.

In (1) the ground-state total energies of the clusters with one extra electron in the LUMO (C_{60}^-) and with a hole in the HOMO (C_{60}^+) are needed. To determine E_{N+1}^{TOT} a calculation for the anionic C_{60}^- has been performed with the added electron in the t_{1u} state. This configuration is achieved by adding 1/3 of the electron charge to each state of the 3-fold degenerate first empty state of the cluster. An analogous calculation has been performed to determine the total energy of the cationic C_{60}^+: the h_u state has been filled with nine electrons distributed uniformly in the 5-fold degenerate states. For both cases the corresponding eigenvalues have been also considered, since an alternative expression for the energy gap reads[3]

$$E_{gap} \simeq \epsilon_{N+1}^{N+1} - \epsilon_N^N, \tag{2}$$

where ϵ_M^N is the M-th eigenvalue of a N-electrons system. Note that while (1) is exact, (2) is an approximate expression valid in the limit $N \to \infty$. In Tab.(I) are reported the QP gap values obtained from (1) and (2). In both cases a HOMO-LUMO gap larger than the DFT-LDA one more than 3eV has been obtained, the 2nd being slightly larger than the first one. The calculated increase of the DFT-LDA energy difference between the h_u (occupied) and t_{1u}(empty) molecule states could be with these values slightly overestimated. This seems to be especially valid when comparing with results for solid C_{60} films combined from photoemission and inverse photoemission. The comparison of our data with experimental ones, namely the difference between vertical ionization potential(I) and electron affinity(A), to be taken with their error bars, is satisfactory. To summarize for C_{60}, the DFT-LDA gap underestimates the experimental one (taken as the average value in the range reported in Tab.(I)) by more than 64% while that obtained from (1) reproduces it within 5%. The single-particle spectrum of the cluster can be been obtained from the DFT-LDA one by applying the *scissor operator* rule with a 3.42eV calculated gap correction. This approximate rule to correct the *band-gap problem*, consisting in rigidly moving up in energy the empty states mantaining fixed the filled ones , has been widely used for QP spectrum calculations of bulk semiconductors. Its approximate validity has been recently demonstrated also in the sodium tetramer case.[7]

Using the same approach we can give an estimate of the first excitonic transition energy using the expression

$$E_{2p} = E_N^{TOT} - E_N^{TOT}(*), \tag{3}$$

where, on the right-hand side, the difference between the ground-state energy of the neutral cluster and that of the excited neutral one appear. E_{2p} corresponds to the energy difference between the first two peaks of the photodetachment spectrum of charged clusters (see Ref.[6] and references therein). E_{2p} is indeed the energy gap between the ground and the first excited state of the two-particle spectrum ($^1A_g \to ^1 T_{1g}$) and should not be confused with the single-particle HOMO-LUMO gap of (1) and (2). The value of E_{2p} appearing in (3) has been evaluated within the DFT-LDA scheme outlined previously in this work , by placing an electron in the t_{1u} state and a hole in the h_u; a value of 1.82eV results, in good agreement with the experimental ones of 1.9eV, and 1.81eV. The large energy difference between the single-particle gap ($\sim 5eV$) and the two-particle one ($\sim 2eV$) can be ascribed

HOMO-LUMO gap Na$_4$	present work	previous results
DFT-LDA	0.67	0.55
HF		3.4
GW		3.0
I-A		3.43-3.63
E(1)	3.98	
E(2)	4.20	

Table II: HOMO-LUMO gap values in eV of neutral sodium tetramer Na$_4$ as determined in this paper by standard DFT-LDA calculations and according to Eqs.(1) [E(1)] and (2) [E(2)] are compared with previous theoretical results. An estimate for the experimental difference between vertical ionization potential and electron affinity is also reported.

to strong excitonic effects (binding energies $\sim 3eV$) occouring in the cluster excited with an electron-hole pair.

II) The Sodium Tetramer Na$_4$

For the sodium tetramer the same calculation scheme has been followed as for the fullerene cluster. In this case a full or *ab initio*(so stressing that the full dielectric function and the one-electron Green function have been calculated from eigenvalues and eigenfunctions of the system in constrast to simplified approaches) GW result has been recently presented for the quasiparticle spectrum of the small sodium cluster.[7] A HOMO-LUMO gap of 3eV has been obtained with a self-energy correction of more than 2.5eV. In Tab.(II) the LDA gap as well as the QP ΔSCF gaps from (1) and (2) have been compared with previous results. We have assumed the same rhombic planar geometry, with D_{2h} symmetry obtained by all previous authors, and we have minimized cluster total energy in order to obtain equilibrium parameters. All-electron DVM calculations have been performed for Na$_4$ with a basis set made up of atomic occupied orbitals plus 3p empty orbitals, in order to assure higher variational flexibility in the bond region. A 300-point grid has been tested to give full numerical stability to our matrix elements. At equilibrium we got the values 5.757 a.u. and 11.901 a.u. for the two axes and of 0.67 eV for the LDA HOMO-LUMO gap. In Tab.(II) we have also reported the HF and GW (see Ref.[7]) results for the gap as well as the experimental data.

It is clear from the table that both our results overestimate the experimental and the GW ones. While DFT-LDA result underestimates the experiment of about 81% (and the GW one of 78%), by using (1) the overestimate turns out to be 13% with respect to the experiment and about 30% with respect to the GW one. This larger overestimate with respect to the fullerene case can be ascribed to the well-known error in the long range behaviour of the DFT-LDA exchange-correlation energy functional. We expect that this error is less important in systems where the orbitals are more delocalized(in C$_{60}$ than in Na$_4$). In other words, the electronic relaxation effects taken into account via (1) give larger contribution to the gap in the more localized system; therefore a larger LDA error is expected in Na$_4$ than in the fullerene molecule. Moreover, in Tab(II) the difference between E(1) and E(2) is almost twice as much as in the fullerene case (Tab.(I)). This is actually the expected trend followed by the approximated expression for the gap (2) as the number of particles increases.[3] Finally, also in the Na$_4$ case QP spectrum may be easily obtainable

from the DFT-LDA results using the *scissor operator* rule, as confirmed by more exact GW calculations[7]. Its comparison with the optical spectrum indicates also in this system the appearence of strong excitonic effects.In this case the energy E(3) is 0.89eV, thus in good accordance with the value of 1.1eV obtained subtracting to the GW HOMO-LUMO gap the optical absorption onset.[7]

III) CONCLUSIONS

In summary an efficient method to the calculation of the HOMO-LUMO gap and electron single particle spectrum of clusters has been applied to the fullerene and to the sodium tetramer cases obtaining good agreement with available experimental and theoretical data, the results for the sodium cluster being less satisfactory due to the LDA limits. This method appears of order of magnitudes more efficient than schemes based on different hamiltonians (e.g. *ab initio* methods), since its computations have been performed on vax-class microcomputers. Therefore it may be very convenient to takle the problem of the QP spectra calculations of large-dimension clusters prior to their full solution by *ab initio* methods.

REFERENCES

1. R.O. Jones and O.Gunnarson, Rev. Mod. Phys. **61**, 689 (1989)

2. M.S. Hybertsen S.G. Louie, Phys. Rev.Lett. **55**, 1418 (1985); Phys.Rev.B **34**, 5390

3. R.W. Godby, M. Schlüter, L.J. Sham, Phys. Rev. **37**, 10159 (1988)

4. D.E. Ellis and G.S. Painter, Phys. Rev. B **2**, 2887 (1970); Yang Jinlong, Wang Kelin, F.Casula, G. Mula, Phys. Rev. **47**, 4025 (1993) ; Yang Jinlong, F. Toigo, and Wang Kelin, Phys. Rev. **50**, 7915 (1994); and references therein

5. G. Cappellini,F. Bechstedt,A. Bosin., F. Casula, Phys. Sta. Sol.(b)**189**, 153 (1995)

6. G. Cappellini,F. Casula, Jinlong Yang, F. Bechstedt, Phys. Rev. **56**, 3628 (1997)

7. G.Onida, L. Reining, R.W. Godby, R. Del Sole, W. Andreoni Phys. Rev. Lett. **75**, 818 (1995)

Part II

Semi–Empirical Tight Binding

AB INITIO CALCULATION OF TIGHT-BINDING PARAMETERS

A. K. MCMAHAN, J. E. KLEPEIS
Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94551

ABSTRACT

We calculate *ab initio* values of tight-binding parameters for the *f*-electron metal Ce and various phases of Si, from local-density functional one-electron Hamiltonian and overlap matrix elements. Our approach allows us to unambiguously test the validity of the common minimal basis and two-center approximations as well as to determine the degree of transferability of both nonorthogonal and orthogonal hopping parameters in the cases considered.

INTRODUCTION

The success of local-density functional calculated total energies and forces and the synthesis of these techniques with molecular dynamics has opened up major new possibilities for atomistic simulation [1, 2]. The expense of such *ab initio* molecular dynamics calculations, however, remains a drawback which has focused attention on the use of cheaper electronic calculations as in tight-binding molecular dynamics [3]. The tight-binding total-energy representations required for this application are typically developed by fitting large *ab initio* generated data bases of band structures, total energies, and possibly forces over the range of stoichiometry, coordination, and atomic volume which are of interest [4]–[10]. Unfortunately the nonlinear optimization required in such fits becomes more and more difficult as the number of chemical species increases, and the time required to generate the data bases can also become prohibitive. Ideally one would like to automate the process, perhaps with *ab initio* calculation of the representation itself. While this is a challenging goal, it may be useful to begin with an examination of what our local-density functional methods are actually doing from a tight-binding perspective. This approach enables direct tests of the accuracy and transferability of the resulting tight-binding parameters while at the same time providing insight into the general characteristics required for a robust parametrization.

In a recent paper [11], we have shown that tight-binding parameters may be directly and simply calculated from the k-dependent one-electron Hamiltonian and overlap matrix elements generated by any band structure method which is formulated in or can be projected onto a localized basis. After briefly reviewing these theoretical concepts and the FP-LMTO method [12] which we use to calculate these matrix elements, we extend such calculations here to include an *f*-electron metal, Ce. By a comparison of the *ab initio* bands and those reconstructed from the calculated tight-binding parameters we assess two major approximations customary in tight-binding: use of a minimal basis and the two-center expansion. We then report new results for a range of coordinations in Si, and discuss the issue of transferability in terms of the *ab initio* calculated nonorthogonal parameters. Finally we turn to the impact of orthogonalization on transferability, again for the case of Si, and then provide a summary.

THEORETICAL METHODS

Tight-binding Parameters

To facilitate subsequent discussion, it is useful to briefly review the formal relations between matrix elements of the one-electron Hamiltonian, H, or overlap, O, operators and Slater-Koster tight binding parameters. For simplicity, we consider only one atom per primitive cell, as more general expressions are given elsewhere [11]. In particular, let $|\mathbf{R}\ell m\rangle$ designate an orbital of angular and magnetic quantum numbers ℓ and m, respectively, at the lattice site $\mathbf{R}$. Direct and k-space matrix elements of H are related as usual by

$$\langle 0\ell m|H|\mathbf{R}\ell'm'\rangle = \frac{1}{N}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}}\, H^{\mathbf{k}}_{\ell m,\ell'm'}, \tag{1}$$

where the former are real as we consider m to represent the customary real linear combinations of orbitals for each ℓ [13, 14] .

The intersite ($R\neq 0$) matrix elements may be approximated by a two-center expansion in terms of Slater-Koster hopping parameters, $t_{\ell\ell'\mu}(R)$.

$$\langle 0\ell m|H|\mathbf{R}\ell'm'\rangle \approx \sum_{\mu} g_{\mu}(\ell m, \ell'm', \hat{\mathbf{R}})\, t_{\ell\ell'\mu}(R) \tag{2}$$

Here, $g_{\mu}(\ell m, \ell'm', \hat{\mathbf{R}})$ are (real) geometric factors [13, 14], μ is the magnetic quantum number about the bond axis $\mathbf{R}$, $\mu = \sigma, \pi, \cdots, \min(\ell, \ell')$, and $R \equiv |\mathbf{R}|$, $\hat{\mathbf{R}} \equiv \mathbf{R}/R$. Orthogonality relations [11] among the g_{μ} give

$$t_{\ell\ell'\mu}(R) = (2 - \delta_{\mu\sigma})^{-1} \sum_{m,m'} g_{\mu}(\ell m, \ell'm', \hat{\mathbf{R}})\, \langle 0\ell m|H|\mathbf{R}\ell'm'\rangle, \tag{3}$$

which in conjunction with Eq.(1), uniquely defines the Slater-Koster hopping parameters in terms of the k-dependent one-electron Hamiltonian matrix elements. Hopping parameters, $s_{\ell\ell'\mu}(R)$, are similarly defined from the overlap matrix elements.

The essential nature of the two-center approximation in Eq.(2) is the use of the Slater-Koster geometric functions, $g_{\mu}(\ell m, \ell'm', \hat{\mathbf{R}})$, as a matrix basis for the m,m'-dependence of the $\langle 0\ell m|H|\mathbf{R}\ell'm'\rangle$. Equation (2) is approximate since this basis is in general incomplete. The $t_{\ell\ell'\mu}$ are nonetheless precisely defined as the projections of $\langle 0\ell m|H|\mathbf{R}\ell'm'\rangle$ onto the existing matrix basis functions. Three-center contributions to the matrix elements arising from one-electron potential terms in H close to the bond axis, $\mathbf{R}$, are likely also to have significant projections onto these basis functions. "Non-two-center" effects is therefore perhaps a better description of the inadequacies of Eq.(2) than is three-center terms. Conversely, the extent to which three-center terms are absorbed into the effective $t_{\ell\ell'\mu}$ is a source of non transferability as will be discussed subsequently.

The intrasite matrix elements, $\langle 0\ell m|H|0\ell'm'\rangle$, define site energy, ε_{ℓ}, and Slater-Koster crystal field parameters, $\chi_{\ell\ell'\mu}(R)$.

$$\langle 0\ell m|H|0\ell'm'\rangle \approx \varepsilon_{\ell}\delta_{\ell\ell'}\delta_{mm'} + \sum_{\mathbf{R},\mu}^{R\neq 0} g_{\mu}(\ell m, \ell'm', \hat{\mathbf{R}})\, \chi_{\ell\ell'\mu}(R) \tag{4}$$

The latter arise from two-center expansion of one-electron potential terms in H at sites $\mathbf{R}$. In contrast to the hopping parameters, the ε_{ℓ} and $\chi_{\ell\ell'\mu}(R)$ are underdetermined by the

matrix elements of a *single* structure. The intrasite matrix elements behave as if they were characterized by generally a small number of *effective* site energies, $\varepsilon_\ell^{\text{eff}}$, and crystal field interactions, $\chi_{\ell\ell'\mu}^{\text{eff}}(R)$, which can be determined. The former are given by

$$\varepsilon_\ell^{eff} = \frac{1}{2\ell+1}\sum_m \langle 0\ell m|H|0\ell m\rangle, \tag{5}$$

$$= \varepsilon_\ell + \frac{1}{2\ell+1}\sum_{\mathbf{R},\mu}^{R\neq 0}(2-\delta_{\mu\sigma})\chi_{\ell\ell\mu}(R), \tag{6}$$

showing that site energies extracted from either one-electron matrix elements or band structure are intrinsically dependent on local environment. The similarity of Eq.(6) to a pair potential sum, moreover, provides justification [11] for using environmentally dependent site energies in place of interatomic potentials in tight-binding total-energy representations [7]. The remaining effective crystal field parameters are discussed elsewhere [11].

FP-LMTO Calculations

We use the FP-LMTO method [12] to obtain the minimal basis matrix elements from which the tight binding parameters are calculated according to the above expressions. Details are given elsewhere [11]. In brief, the FP-LMTO method is first used in its most rigorous multiple-κ form to obtain self-consistent one electron potentials and accurate band structures for test comparisons. A linked or contracted minimal basis is then chosen by minimization of the occupied one-electron eigenvalue sum using the fixed self-consistent potentials. The variables optimized in this way include the relative coefficients of the two augmented spherical Hankel functions which form each minimal basis function. The decay energies, $-\kappa^2$, of the Hankel function tails correspond to $\kappa_1^2 = 0.5$ Ry with κ_2^2 also optimized in the range 1–1.5 Ry. The tight binding parameters are then calculated from the resultant minimal basis Hamiltonian and overlap matrices. When discussing issues of transferability between different structures, as here for Si, we also first perform a unitary transformation which "rotates" all Hamiltonian and overlap matrices so as to guarantee a perfectly transferable, two-center overlap while preserving the band structure. This allows all issues of non-two-center effects and transferability to be addressed solely in terms of the Hamiltonian hopping parameters. All muffin-tin radii in both multiple-κ and minimal basis calculations were taken to be nearly (reduced by ~1%) touching, unless otherwise stated.

APPROXIMATIONS

Tight binding representations of the one electron Hamiltonian typically make two approximations: (1) use of a minimal basis, and (2) the two-center approximation. We examine these approximations separately here for the case of fcc Ce at the experimental volume of the α phase.

Figure 1(a) compares our best multiple-κ band structure (solid lines) to the linked minimal basis representation (dots). The former calculations included three (κ^2 =0.01, 1, 2.3 Ry) augmented spherical Hankel basis functions of each ℓm type for s and p states, two (κ^2 =0.01, 1 Ry) for d states, and one (κ^2 =0.01 Ry) for f states. The dimension of the Hamiltonian and overlap matrices was therefore $29 = 3 \times 1 + 3 \times 3 + 2 \times 5 + 1 \times 7$ at each **k** point. Our linked minimal basis treatment used two augmented spherical
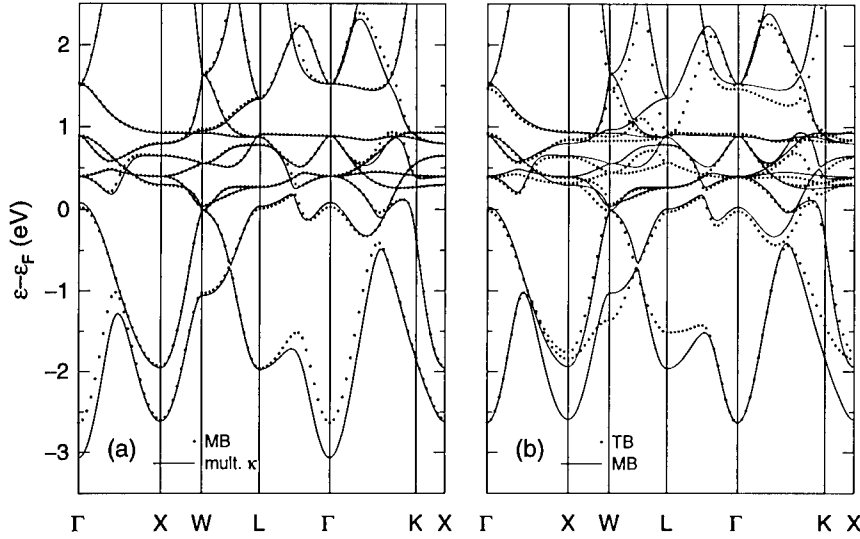
Figure 1: Band structure for fcc Ce: (a) comparison of our best multiple-κ bands (lines) to the results from the minimal basis (dots), and (b) comparison of the minimal basis bands (lines) to the tight-binding bands reconstructed from the tight-binding parameters calculated from the minimal basis matrix elements (dots).

Hankel functions (κ^2 =0.5, 1.2 Ry) to represent each atomic orbital; however, the relative coefficients were fixed so that the linear combination for each ℓm served as a single atomic orbital, leading to Hamiltonian and overlap matrices of dimension 16. The important distinction between multiple-κ and linked calculations is that the relative coefficients of different augmented spherical Hankel functions corresponding to the same ℓm may vary with k-vector for different eigenfunctions in the first case, whereas they are fixed and k-independent in the latter. These coefficients come out of solving the one-electron eigenvalue problem in the first case, while they were obtained by optimization of the occupied one-electron eigenvalue sum, as noted above, in the latter case.

The agreement between the two sets of bands in Fig. 1(a) is on the whole quite good. The rms differences are 0.17 and 0.02 eV for the two lowest bands, and 0.02±0.01 eV for the seven predominantly f bands between 0.2-1 eV, as determined at 145 points throughout the Brillouin zone. The only significant discrepancy is in the lowest band, primarily at and nearby the pure-s Γ point. We believe this is due to the minimal basis orbital being too localized, given analogous behavior for other materials where we have sometimes found unusually slow convergence at s-like Γ points as a function of shell cut-off in our *ab initio* tight-binding representations. Moreover, since s-s contributions from different directions and shells should contribute constructively at the Γ point, a symptom of a too-localized s basis orbital should be a too-narrow s band with a bottom which is too high in energy as is seen. While we would expect smaller values of κ^2 to remedy this problem in the minimal basis treatment, this would also lead to longer ranged tight-binding representations.

Figure 1(b) compares the minimal basis band structure for fcc Ce (solid lines) to the

tight binding bands (dots) reconstructed using hopping parameters given by Eq.(3) and its overlap analog, the effective site energies given by Eq.(5), and four effective near-neighbor crystal field interactions (one pf, one dd, and two ff) obtained as discussed in [11]. The cut-off for the tight binding parameters was $d_{max}/a = 2.25$, where a is the lattice constant. We verified convergence with d_{max} by obtaining essentially identical bands using $d_{max}/a = 2.99$.

The rms differences between the two lowest bands in Fig. 1(b) are 0.24 and 0.10 eV, respectively, and 0.08 ± 0.04 eV for the seven predominantly f bands. The most serious discrepancies occur for the lowest band, at and near the mostly d-like X (sd hybrid), W (pdf), and L points (sd). All differences between the two sets of bands in Fig. 1(b) are due entirely to non-two-center effects, i.e., to m, m' dependence of the matrix elements in Eq.(1) which can not be spanned by the two-center geometric functions. Since such dependence is likely to arise from three-center contributions to the matrix elements, and these should become less important for more localized bases, one would expect better agreement between the minimal basis and tight-binding bands for more localized bases. We find this to be the case in calculations where we have taken the minimal basis orbitals to be just the $\kappa^2 = 0.5$ Ry augmented spherical Hankel functions. The near-neighbor overlap hopping parameters for this basis are smaller by 24, 47, and $\sim 70\%$ for the $ss\sigma$, $sd\sigma$, and $dd\mu$ parameters, respectively, compared to the linked basis. The corresponding rms agreement for the lowest band between minimal basis and tight binding results is improved for this basis by a factor of 4, to 0.06 eV, over the result using the linked minimum basis. Unfortunately, while there is improved agreement between the minimal basis and tight binding bands, both are in poor agreement with our rigorous multiple-κ bands.

The rms agreement between minimal basis and tight binding $4f$ bands in Fig. 1(b) is about 10% of the $4f$ band width. One might have thought that the $4f$ orbitals were already sufficiently localized even in the linked basis to be immune from non-two-center effects. Note, however, that the number of two-center geometric matrices goes up only as ℓ while the size of the m, m' matrices which they are asked to represent increases by $(2\ell + 1)^2$. It would appear that non-two-center effects become worse with increasing ℓ.

TRANSFERABILITY

In previous *ab initio* calculations of the hopping parameters for the silicon-boron system, we obtained relatively transferable parameters in so far as chemical and atomic-volume variation were concerned, however, there were indications of significant coordination dependence in the parameters [11]. We explore the coordination issue further here for the face centered cubic (fcc), body centered cubic (bcc), simple cubic (sc), and cubic diamond (cd) phases of Si, i.e., coordinations of 12, 8, 6, and 4, respectively.

Figure 2 shows the ratios $-t_{\ell\ell'\mu}(d)/s_{\ell\ell'\mu}(d)$ for the sp subset of our calculated parameters for the four phases of Si. The calculated values are shown as the data points which appear in triples for each neighbor distance, corresponding to the theoretical equilibrium volume, V_0, and $(1 \pm 0.2)V_0$. The lines connecting these points pass through the first neighbors for all three volumes, then the second, and so on. These parameters correspond to the rotated matrices described above, so that the $s_{\ell\ell'\mu}$ are perfectly transferable by construction. A rigid shift of any band structure by $\Delta\varepsilon$ will change the corresponding $t_{\ell\ell'\mu} \rightarrow t_{\ell\ell'\mu} + \Delta\varepsilon\, s_{\ell\ell'\mu}$ in the case of a nonorthogonal basis. We have therefore plotted the $-t_{\ell\ell'\mu}/s_{\ell\ell'\mu}$ ratios in Fig. 2 on the same energy scale, so that such a rigid band shift for a given structure may
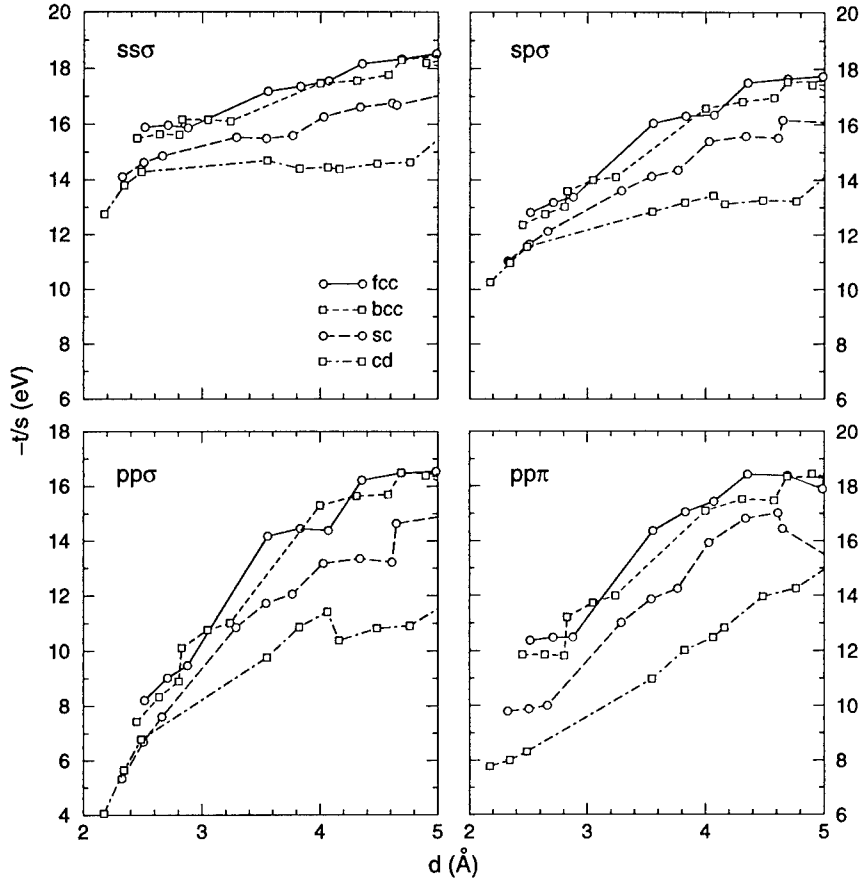
Figure 2: Negative ratios of the Hamiltonian and overlap hopping parameters, $-t_{\ell\ell'\mu}/s_{\ell\ell'\mu}$, for Si structures.

be envisioned by moving all four $\ell\ell'\mu$ curves for that structure up or down by the *same* amount. No such shifts have been performed; however, the results in Fig. 2 are as obtained directly from the minimal basis matrix elements. The full set of calculated *spd* parameters reproduce our best multiple-κ band structures to within an rms of 0.1–0.2 eV for states up to 4 eV above the Fermi level, an energy range of 15–20 eV. Note that the actual magnitudes of the t's change by factors of ~20 or more over the d range covered in Fig. 2, and that one could draw smooth curves through the calculated $t_{\ell\ell'\mu}(d)$ for all volumes and neighbors of a given structure which would follow this variation to within about 5% or so.

The essential conclusions to be drawn from Fig. 2 are that the high coordination fcc and bcc parameters track each other fairly well while the lower coordination structures have systematically smaller t's. Even if rigid band shifts were used to align the near-neighbor parameters more closely for all structures, the latter comment would still apply to the 2nd

and 3rd neighbor values with spreads of 10–20% and 15–35%, respectively [15]. In the remainder of this section, we discuss these conclusions to see if they are consistent with intuitive expectations.

Suppose that the self-consistent one-electron Hamiltonian can be approximated by

$$H = \frac{-\hbar^2}{2m}\nabla^2 + \sum_{\mathbf{R}} v_{\mathbf{R}}, \tag{7}$$

where the $v_{\mathbf{R}}(\mathbf{r})$ are spherically symmetric about $\mathbf{R}$ and

$$(\frac{-\hbar^2}{2m}\nabla^2 + v_{\mathbf{R}})|\mathbf{R}\ell m\rangle = \varepsilon_\ell|\mathbf{R}\ell m\rangle. \tag{8}$$

It then follows from Eq.(3) and the g_μ orthogonality relations that

$$
\begin{aligned}
t_{\ell\ell'\mu}(R) &= \frac{1}{2}(\varepsilon_\ell + \varepsilon_{\ell'})\, s_{\ell\ell'\mu}(R) + \Delta t_{\ell\ell'\mu}(R) \\
&\quad + (2 - \delta_{\mu\sigma})^{-1} \sum_{\mathbf{R}'}^{\neq 0,\mathbf{R}} \sum_{m,m'} g_\mu(\ell m, \ell'm', \hat{\mathbf{R}})\,\langle 0\ell m|v_{\mathbf{R}'}|\mathbf{R}\ell'm'\rangle,
\end{aligned} \tag{9}
$$

where $\Delta t_{\ell\ell'\mu}(R)$ comes from the two-center expansion of $\langle 0\ell m|\frac{1}{2}(v_0 + v_{\mathbf{R}})|\mathbf{R}\ell'm'\rangle$. While the $v_{\mathbf{R}}$ may of course be structure dependent, a more obvious source of environmental dependence of the $t_{\ell\ell'\mu}$ will come from the three-center matrix elements in the last term of Eq.(9). This expression is particularly simple for the $ss\sigma$ case where $g_\sigma(s, s, \hat{\mathbf{R}}) = 1$. Since one might expect the $v_{\mathbf{R}}$ to be negative, these three-center contributions to $t_{ss\sigma}$ should then all make positive contributions to the $-t_{ss\sigma}/s_{ss\sigma}$ ratios plotted in Fig. 2. An examination of the fcc, bcc, cd, and sc structures shows that the closer packed phases tend to have more atoms closer to the shorter bond axes $\mathbf{R}$ than do the lower coordination structures, which is consistent at least on geometric grounds with the $-t_{ss\sigma}/s_{ss\sigma}$ ordering in Fig. 2.

Additional insight is provided by the canonical band theory of Andersen and coworkers [16]. For pure ℓ-bands, one may write the second-order one-electron eigenvalue problem in the form

$$[C_\ell + (\Delta_\ell - \gamma_\ell C_\ell)\mathbf{S}_{\ell\ell}]x = \lambda(1 - \gamma_\ell\mathbf{S}_{\ell\ell})x, \tag{10}$$

where λ is the eigenvalue and $\mathbf{S}_{\ell\ell} = [S_{\mathbf{R}\ell m, \mathbf{R}'\ell m'}]$ is the bare canonical structure constant matrix. Material dependent potential parameters are the band center, C_ℓ, the band width or inverse mass, Δ_ℓ, and a band distortion or nonorthogonality parameter, γ_ℓ. Given the definition of the canonical structure constants [16], it can be shown that Eq. (10) is formally equivalent to the nonorthogonal tight-binding problem

$$
\begin{aligned}
\varepsilon_\ell &= C_\ell, \tag{11} \\
s_{\ell\ell\mu}(d) &= \gamma_\ell\, f_{\ell\ell\mu}\, 2(2\ell + 1)(R_{\mathrm{WS}}/d)^{2\ell+1}, \tag{12} \\
t_{\ell\ell\mu}(d) - \varepsilon_\ell s_{\ell\ell\mu}(d) &= -\Delta_\ell\, f_{\ell\ell\mu}\, 2(2\ell + 1)(R_{\mathrm{WS}}/d)^{2\ell+1}, \tag{13}
\end{aligned}
$$

where R_{WS} is the Wigner-Seitz radius, $f_{ss\sigma} = 1$, $f_{pp(\sigma,\pi)} = -2, 1$, and $f_{dd(\sigma,\pi,\delta)} = 6, -4, 1$.

One of the strengths of the LMTO-ASA method [16] has been its separation of material dependence (potential parameters) from structure dependence (structure constants). The tight-binding representation given by Eqs. (11–13) is explicitly transferable to the extent that the C_ℓ, γ_ℓ, and Δ_ℓ are independent of structure, which is known to be the case to a high

degree for the relatively close packed structures for which the LMTO-ASA method is most accurate. Moreover, there has been considerable success in the calculation of structural energy differences for relatively close packed materials using canonical band theory [17] as well as by the LMTO force-relation technique [18] which both embody these same ideas of using fixed potential parameters which we equate to transferable representations. All of this experience, as well as the success of new tight-binding total energy representations for the transition metals [7], suggest a rather high degree of transferability in the close packed limit, in agreement with Fig. 2.

ORTHOGONAL BASES

Here we consider the suitability of orthogonal bases by performing Löwdin transformations [19], $\mathbf{H}^{\mathrm{or}} = \mathbf{O}^{-1/2}\mathbf{H}\mathbf{O}^{-1/2}$, to create (k-dependent) Hamiltonian matrix elements, $\mathbf{H}^{\mathrm{or}}$, in an orthogonal basis. Unfortunately, we find that the Löwdin transformation amplifies non-two-center effects, as may be seen from the expansion in $\mathbf{\Delta O} \equiv \mathbf{O} - \mathbf{1}$.

$$\mathbf{H}^{\mathrm{or}} = \mathbf{H} - \frac{1}{2}(\mathbf{\Delta O} \cdot \mathbf{H} + \mathbf{H} \cdot \mathbf{\Delta O}) + \cdots \qquad (14)$$

Terms which are of first and higher order in $\mathbf{\Delta O}$ involve products of the geometric matrices which, due to the incompleteness of the $\mathbf{g}_\mu(\ell, \ell', \hat{\mathbf{R}}) \equiv [g_\mu(\ell m, \ell'm', \hat{\mathbf{R}})]$, add non-two-center contributions to the $\mathbf{H}^{\mathrm{or}}$ matrix elements. We consistently find degraded agreement between the orthogonal Hamiltonian matrix elements, $\langle 0\ell m|H^{\mathrm{or}}|\mathbf{R}\ell m'\rangle$ and their tight binding reconstructions, as well as between the corresponding band structures, in comparison to the nonorthogonal case. These non-two-center effects diminish for more localized bases. In order to focus solely on the effects of orthogonalization in this section, we use a basis which is sufficiently localized that the minimal basis, nonorthogonal tight binding, and orthogonal tight binding band structures are all three essentially identical. This agreement comes at the cost, however, of poor representation of the accurate multiple-κ band structure.

Figure 3 shows calculations of nonorthogonal, t, and orthogonal, t^{or}, hopping parameters for the fcc and cd phases of Si, with appropriate sign changes as required to plot positive numbers. As in Fig. 2, the calculated parameters appear in triples corresponding to V_0 and $(1 \pm 0.2)V_0$. They are connected by lines (solid, fcc; dashed, cd) for visual convenience, and the nonorthogonal parameters have been multiplied by three ("$3t$") to separate the data. The minimal basis chosen corresponded to a single augmented spherical Hankel function ($\kappa^2 = 0.5$ Ry) for each atomic orbital. In addition, we used smaller, fixed muffin-tin radii ($R_{\mathrm{MT}} = 1.7$ Bohr) for all atoms in order to switch sooner from the numerically generated radial wave functions to the faster decaying Hankel function tails. For comparison, touching muffin-tins in $V/V_0 = 0.8$ cd Si would have $R_{\mathrm{MT}} = 2.05$ Bohr. The present nonorthogonal-case overlap parameters, $s_{\ell\ell'\mu}(d)$, are already so consistent among the two structures and three volumes, that there was no need to perform unitary transformations to guarantee this transferability. The sp subset of these $s_{\ell\ell'\mu}(d)$ are smaller by factors of 1.6–3.8 at the near-neighbor distance, and by factors of 2.2–6.9 at the third neighbor distance, in comparison with our optimum linked-basis parameters used in Fig. 2, which provided good agreement with the multiple-κ bands.

The first feature to note in Fig. 3 is that the nonorthogonal t's are in relatively close agreement between the two structures and for all three volumes. While the fcc near-neighbor $t_{\ell\ell'\sigma}$'s appear to be slightly lower than their cd counterparts, recall that a rigid-
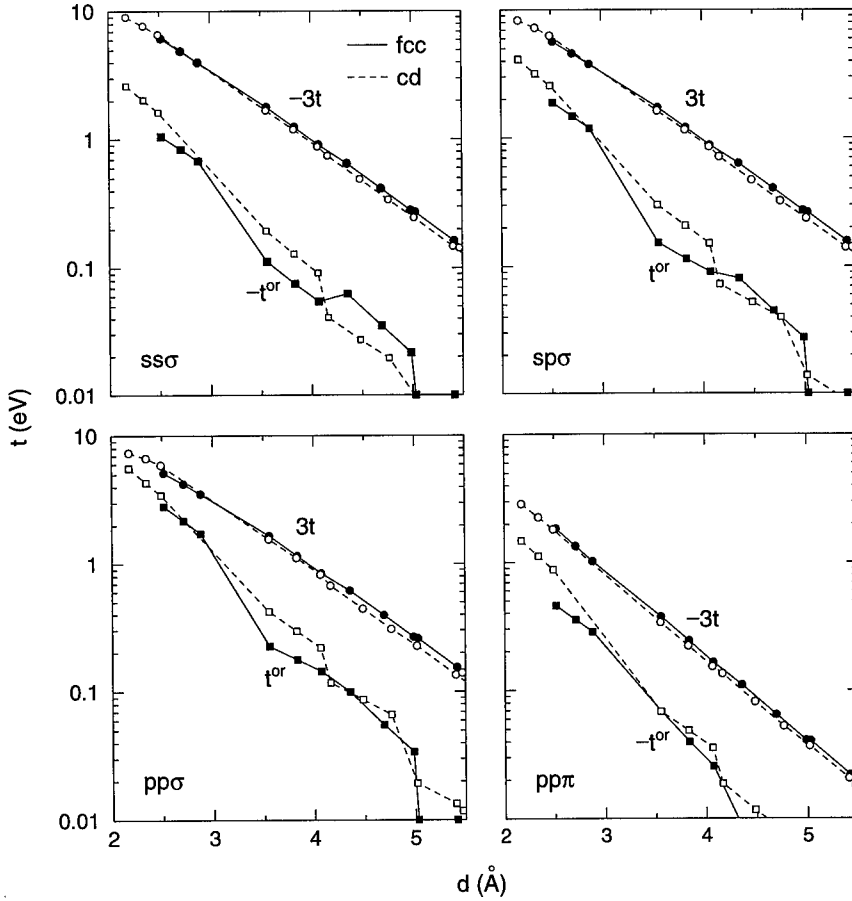
Figure 3: Nonorthogonal hopping parameters multiplied by three (circles, $3t$) and orthogonal hopping parameters (squares, t^{or}) for fcc (filled symbols, solid lines) and cd Si (open symbols, dashed lines). Some fcc t^{or} for $d > 5$ Å have changed sign, which is indicated by plotting these points on the bottom axes of the three $\ell\ell'\sigma$ panels.

band shift could have achieved the same result for Fig. 2. More significant is the slight rise here with increasing d of the fcc t's relative to the cd values, a trend which was fully evident in Fig. 2. These are 10–20% effects, whereas the orthogonal basis hopping parameters, t^{or}, in Fig. 3 are clearly much less transferable. The $V/V_0 = 1.2$ cd and $V/V_0 = 0.8$ fcc near-neighbor distances are nearly the same, yet the cd t^{or} magnitudes here are larger by factors of 1.2–1.9 than the corresponding fcc values. Similarly, the second neighbor distances are fairly close, and here the cd magnitudes are a factor of ~1.7 larger than the fcc values.

We have analyzed the effect of orthogonalization using the first order terms in $\mathbf{\Delta O}$ shown in Eq.(14). To verify that this would reasonable, we first numerically calculated the

$t^{\mathrm{or}}(d)$ over the range $0 \le \eta \le 1$, corresponding to the Löwdin orthogonalization of $\mathbf{H}$ using a scaled overlap, $1 + \eta(\mathbf{O}-\mathbf{1})$. The usual nonorthogonal, t, and orthogonal, t^{or}, parameters correspond to the $\eta = 0$ and 1 limits, respectively. We generally found that the first order corrections, $(dt^{\mathrm{or}}/d\eta)|_{\eta=0}$, to the t's were of the same sign although significantly larger (factors of 2 or so) than the exact corrections, $t^{\mathrm{or}}-t$. Given that the signs of the first order corrections tend to be right, we then used Eq.(14) to obtain analytic expressions for these terms for further investigation [20]. In the absence of crystal field, Eq.(14) implies

$$
\begin{aligned}
t^{\mathrm{or}}_{\ell\ell'\mu}(R) = {}& t_{\ell\ell'\mu}(R) - \frac{1}{2}(\varepsilon_\ell + \varepsilon_{\ell'})\, s_{\ell\ell'\mu}(R) \\
& - \frac{1}{2(2-\delta_{\mu\sigma})} \sum_{\mathbf{R'}}^{\ne 0,\mathbf{R}} \sum_{\lambda,\mu_1,\mu_2} \mathbf{g}_\mu(\ell,\ell',\hat{\mathbf{R}}) \cdot \mathbf{g}_{\mu_1}(\ell,\lambda,\hat{\mathbf{R}'}) \times \mathbf{g}_{\mu_2}(\lambda,\ell',\widehat{\mathbf{R}-\mathbf{R'}}) \\
& \times [t_{\ell\lambda\mu_1}(R')\, s_{\lambda\ell'\mu_2}(|\mathbf{R}-\mathbf{R'}|) + s_{\ell\lambda\mu_1}(R')\, t_{\lambda\ell'\mu_2}(|\mathbf{R}-\mathbf{R'}|)] + \cdots .
\end{aligned} \tag{15}
$$

Here, $\mathbf{g}_\mu(\ell,\ell',\hat{\mathbf{R}}) \equiv [g_\mu(\ell m, \ell'm', \hat{\mathbf{R}})]$, the $\times$ signifies regular matrix multiplication, and the dot product, $\sum_{m,m'}$, so that the triple product of geometric matrices is a scalar. The bond atoms at 0 and $\mathbf{R}$ together with a neighboring atom at $\mathbf{R}'$ define triangles with angles θ_1 and θ_2 adjacent to the bond, and $\theta_3 = \pi - \theta_1 - \theta_2$ opposite to the bond. The triple product of geometric matrices depends only on these angles, and for the simplest $t^{\mathrm{or}}_{ss\sigma}$ case is 1 and $-\cos\theta_3$ for the $\lambda = s$ and p contributions, respectively. Both make the $t^{\mathrm{or}}_{ss\sigma}$ more positive, i.e., reduce its magnitude, if $\theta_3 < 90°$. More generally we find for an sp basis that most of the contributions tend to reduce the magnitudes of the $t^{\mathrm{or}}_{\ell\ell'\mu}$ if all angles are acute, $0 < \theta_1, \theta_2, \theta_3 < 90°$. It is a characteristic of the more close packed phases that not only are there more neighbors closer to the shorter bond axes, implying larger contributions to Eq.(15), but the associated angles also tend to all be acute so that the t^{or} are generally reduced in magnitude. For more open packed structures, on the other hand, there are fewer nearby neighbors, implying reduced impact of the orthogonalization, and for the cd structure, the smallest triangles have an obtuse angle, so there is greater possibility for cancellation amongst the terms in Eq.(15). These geometric considerations appear consistent with the results in Fig. 3.

SUMMARY

Current interest in tight-binding representations is driven by the need for fast yet still accurate calculation of interatomic forces for use in molecular statics and dynamics calculations. Given the success of local-density functional methods in the calculation of total energies and forces, it makes sense to analyze these methods from a tight-binding perspective both to obtain *ab initio* values of the parameters as well as to gain insight into the validity of approximations and assumptions customarily made in tight-binding treatments.

Two fundamental approximations are made in typical tight-binding representation of one-electron matrices or band structures: use of a minimal basis and the two-center approximation. Although there are some instances where the former approximation could be improved, we find on the whole that the use of a minimal basis is quite satisfactory, as seen both here for fcc Ce and in earlier work [11] for Si/B phases. The two center expansion is by far the more serious approximation. While we find this approximation to be quite good for relatively localized bases, we consistently find those minimal bases which

most accurately reproduce our best multiple-κ bands to be also accompanied by noticeable non-two-center effects. This is particularly evident here in d-like states at the bottom of the Ce conduction band, as well as in the conduction bands of four-fold coordinated Si/B structures [11]. Even the $4f$ states in Ce which should be relatively localized are not immune. In this regard, we have noted that the two-center approximation is likely to become worse with increasing ℓ.

We have also investigated the transferability of our *ab initio* calculated Slater-Koster parameters across a range of coordinations for the case of Si, in both nonorthogonal and orthogonal representations. In agreement with generally held expectations, we find the nonorthogonal parameters to have far better transferability. But even for the nonorthogonal parameters there is residual coordination dependence which probably originates from the three-center terms. It is interesting that three-center terms may well be the source of both of the problems discussed in this work. If one imagines separating each such term into a component orthogonal to and another spanned by the space of two-center geometric matrices, then the former contribute to inadequacies of the two-center approximation, while the latter likely contribute to environmental dependence, i.e., non transferability, of the two-center Slater-Koster parameters. The use of environmentally dependent fitting procedures [9] for the hopping parameter addresses the second but not the first aspect of this compound three-center problem.

It should be acknowledged that tight-binding total-energy representations do not require the one-electron matrix elements themselves, which determine both band structure and eigenvectors, but only the former. The less demanding representational problem of the band structure alone provides more freedom for effective two-center parameters to incorporate what are truly non-two-center effects. Given the fundamental goal of accurate total energies and forces, it may well be that such effective parameters are entirely adequate, and for that matter, that accurate representation of the *ab initio* local density functional bands is not critical. On the other hand, local density functional theory is itself a paradigm for accurate transferable forces, and clear suggestions for improved tight-binding representations arise from its examination. One is to go beyond the two-center approximation by augmenting the current space of two-center geometric functions. A second point is to affirm the use [7] of environmentally dependent site energies as a more flexible format for incorporating the interatomic potential terms ususally incorporated in tight-binding total-energy representations.

ACKNOWLEDGMENTS

REFERENCES

1. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

2. G. Galli, A. Selloni, and R. Car, Chimia **47**, 313 (1993).

3. L. Colombo, Annual Reviews of Computational Physics IV, edited by D. Stauffer (World Scientific, 1996), p. 147.

4. L. Goodwin, A. J. Skinner, and D. G. Pettifor, Europhys. Lett. **9**, 701 (1989).

5. S. Sawada, Vacuum **41**, 612 (1990); M. Kohyama, J. Phys.:Condens. Matter **3**, 2193 (1991).

6. I. Kwon, R. Biswas, C. Z. Wang, K. M. Ho, and C. M. Soukoulis, Phys. Rev. B **49**, 7242 (1994).

7. R. E. Cohen, M. J. Mehl, and D. A. Papaconstantopoulos, Phys. Rev. B **50**, 14694 (1994); M. J. Mehl, D. A Papaconstantopoulos, R. E. Cohen, and M. M. Sigalas, in *Alloy Modeling and Design*, edited by G. Malcolm Stocks (The Minerals, Metals, and Materials Society, Warrendale, Pennsylvania, 1994), p. 25; M. J. Mehl, D. A. Papaconstantopoulos, and R. E. Cohen, Int. J. Thermophys. **16**, 503 (1995).

8. J. L. Mercer, Jr. and M. Y. Chou, Phys. Rev. B **49**, 8506 (1994)

9. M. S. Tang, C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B **53**, 979 (1996).

10. T. J. Lenosky, J. D. Kress, I. Kwon, A. F. Voter, B. Edwards, D. F. Richards, S. Yang, and J. B. Adams, Phys. Rev. B **55**, 1528 (1997).

11. A. K. McMahan and J. E. Klepeis, Phys. Rev. B **56**, 12250 (1997).

12. M. Methfessel, Phys. Rev. **B 38**, 1537 (1988); M. Methfessel, C. O. Rodriguez, and O. K. Andersen, Phys. Rev. **B 40**, 2009 (1989).

13. J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

14. W. A. Harrison, *Electronic Structure and the Properties of Solids*, (Dover, New York, 1989), p. 479.

15. Rigid band shifts were not taken into account in the discussion of Table IV in [11], and would significantly improve the transferability of the "Table II t's" discussed there.

16. O. K. Andersen, O. Jepsen, and D. Glötzel, in *Highlights of Condensed Matter Theory*, edited by F. Bassani, F. Fumi, and M. P. Tosi (North-Holland, Amsterdam, 1985, p. 59; H. L. Skriver, *The LMTO Method* (Springer, Berlin, 1984).

17. J. C. Duthie and D. G. Pettifor, Phys. Rev. Lett. **38**, 564 (1977).

18. For references and an application of the force theorem, see, e.g., A. K. McMahan and J. A. Moriarty, Phys. Rev. B **27**, 3235 (1983).

19. P. O. Löwdin, J. Chem. Phys. **19**, 1396 (1951).

20. The Löwdin expansion has also been discussed by J. Dorantes-Dávila, A. Vega, and G. M. Pastor, Phys. Rev. B **47**, 12995 (1993).

ENVIRONMENT-DEPENDENT TIGHT-BINDING POTENTIAL MODEL

C. Z. WANG, B. C. PAN, M. S. TANG, H. HAAS, M. SIGALAS, G. D. LEE, K. M. HO
Ames Laboratory and Department of Physics, Iowa State University, Ames, IA 50011

ABSTRACT

We have developed a tight-binding model which goes beyond the traditional two-center approximation and allows the tight-binding parameters to scale according to the bonding environment. Our studies show that this environment-dependent tight-binding model improves remarkably the accuracy and transferability of the potential to describe the structures and properties of higher coordinated metallic systems in addition to those of low coordinated covalent systems.

INTRODUCTION

Tight-binding molecular dynamics (TBMD) has recently emerged as a useful method for atomistic simulation study of structural, dynamical and electronic properties of realistic materials [1]. The method includes explicitly quantum mechanical calculations into molecular dynamics and provide a better description of interatomic interactions in the materials in comparison with the classical potentials. At the same time, the method is computationally much faster than the *ab initio* methods because the tight-binding Hamiltonian matrix elements are described by a set of parametrized functions. The method becomes even more attractive and promising due to recent developments in order-N algorithms for electronic calculation and the use of parallel computers [2-5]. Nevertheless, generating an accurate and transferable tight-binding potentials for molecular dynamics simulation of realistic materials is an outstanding challenge. Although enormous effort has been devoted to developing transferable tight-binding potentials in the past years [6-14], the existing tight-binding potentials are successful in the studies of certain properties and materials, but has not always been very transferable to general structures and materials.

We note that one of the major limitations in the transferability of the previous tight-binding potentials is the assumption of two-center approximation for the tight-binding hopping parameters. [15] While the two-center approximation greatly simplify the TB parametrization, neglecting multi-center interactions is justified only when the electrons are well localized in strong covalent bonds. For systems where metallic effects are significant in addition to the covalent bonds, the two-center approximation becomes inadequate. For example, let us consider the tight-binding hopping parameters of silicon in the diamond and in the fcc structures. Each Si atom has four valence electrons which can form strong sp^3 hybrids in the diamond structure. But in the fcc structure, the valence electrons are shared by twelve neighbours. These different bonding situations suggest that the interaction strength in the diamond structure should be stronger than that in the fcc structure even when the interatomic distances are the same in the two structures. It is obvious that the two-center approximation failed to describe these different bonding situations for the two different structures. Another drawback of the two-center approximation is that it can't distinguish between the nearest neighbour and farther neighbour interactions since the interaction strength depends only on the distance between the two atoms.

It seems that a key step to improve the transferability of the tight-binding model is to go

beyond the traditional two-center approximation and allow the tight-binding interactions to be dependent on the bonding environment. In the past several years, the importance of the environment-dependent interactions has received more and more attention. Sawada [8] and subsequently Mercer and Chou [11]. noted that incorporating three-body interactions into the repulsive energy term of the tight-binding potential can give a better description of the energy-volume curves for silicon and germanium in comparison with pure two-center models. Tight-binding models that allow the diagonal matrix elements to be dependent on the environment of the atoms have also been developed for silicon by Mercer and Chou [12] and for transition metals by Cohen et al. [13]. They showed that allowing the diagonal matrix elements of the TB Hamiltonian to be environment dependent can simplify the expression for the repulsive energy term in the tight-binding models. Coordination-dependent hopping integrals between silicon and hydrogen atoms have been considered by Li and Biswas [16] in their Si-H tight-binding model. They found that an environment-dependent Si-H interaction is essential for a correct description of the properties of the interstitial hydrogen atom in the silicon lattice. A tight-binding model that considers the environment dependence of off-diagonal as well as diagonal matrix elements is the effective-medium tight-binding model recently developed by Stokbro et al. [14] In this model, a LMTO-TB Hamiltonian[17] is used to describe the one-electron energy in the total energy model. Application of this model to the study of Si surfaces has produced some interesting results [18].

In this paper, we present a model which allows all tight-binding interaction parameters to be dependent on the bonding environment through two empirical rescaling functions. Based on this model, we have been generating tight-binding potentials for carbon, silicon, germanium, aluminum, niobium, and molybdenum. Our studies show that this new generation of tight-binding model improves remarkably the accuracy and transferability of the potential to describe the structures and properties of higher coordinated metallic systems in addition to those of low coordinated covalent systems.

THE MODEL

In our model, the environment dependence of the tight-binding parameters are modeled through incorporating two empirical scaling functions into the traditional two-center integrals. The first one is a screening function which mimics the electronic screening effects in solid such that the interaction strength between two atoms in the solid becomes weaker if there are intervening atoms located between them. This new approach allows us to distinguish between first and farther neighbor interactions within the same interaction potential without having to specify separate interactions for first and second neighbors. The second function scales the distance between two atoms according to their effective coordination numbers. Longer effective bond lengths are assumed for higher coordinated atoms. The strength of the hopping parameters between atoms i and j are therefore dependent on the coordination number of the atoms: weaker interaction strength for larger-coordinated structures. This new model preserves the two-center form of the tight-binding hopping integral while the multi-center effects are taking into account.

To make the model as sample as possible, the potentials for carbon, silicon, germanium, and aluminum are constructed using an orthogonal sp^3 basis while the d orbitlas are also considered for the transition metal elements. The interaction of two atoms at distance r are described by a function in a form similar to that in the two-center approximation.

The effects of nonorthogonality, multi-center interactions and the flexibility of the basis are taken into account through renormalizing the interaction strength according to the bonding environment.

Specifically, the hopping parameters and the pairwise repulsive potential are expressed as

$$h(r_{ij}) = \alpha_1 R_{ij}^{-\alpha_2} exp[-\alpha_3 R_{ij}^{\alpha_4}](1 - S_{ij}) \tag{1}$$

In this expression, $h(r_{ij})$ denotes the possible types of interatomic hopping parameters $V_{ss\sigma}$, $V_{sp\sigma}$, $V_{pp\sigma}$, $V_{pp\pi}$ etc. and pairwise repulsive potential $\phi(r_{ij})$ between atoms i and j. r_{ij} is the real distance and R_{ij} is a scaled distance between atoms i and j. S_{ij} is a screening function. The parameters α_1, α_2, α_3, and α_4 and parameters for the bond-length scaling function R_{ij} and the screening function S_{ij} can be different for different hopping parameters and pairwise repulsive potential. These parameters will be determined in the fitting procedure. Note that expression Eq.(1) reduces to the traditional two-center form if we set $R_{ij} = r_{ij}$ and $S_{ij}=0$.

We model the screening function as

$$S_{ij} = \frac{exp(\xi_{ij}) - exp(-\xi_{ij})}{exp(\xi_{ij}) + exp(-\xi_{ij})} \tag{2}$$

with

$$\xi_{ij} = \beta_1 \sum_l exp[-\beta_2(\frac{r_{il} + r_{jl}}{r_{ij}})^{\beta_3}] \tag{3}$$

where β_1, β_2, and β_3 are adjustable parameters. Note that ξ_{ij} depends not only on the distance between atoms i and j, but also on the positions of the neighbors of atoms i and j. Maximum screening effect occurs when the atom l is just sitting on the line connecting the atoms i and j (i.e., $r_{il} + r_{lj}$ is minimum). The screening function decays rapidly when the neighboring atoms move away from the line joining atoms i and j. By using the expression of Eq.(2), the screening function varies smoothly from 0 to near 1 as ξ is increased.

The scaling between the real and effective interatomic distance is given by

$$R_{ij} = r_{ij}(1 + \delta_1\Delta + \delta_2\Delta^2 + \delta_3\Delta^3) \tag{4}$$

where $\Delta = \frac{1}{2}[(\frac{n_i-n_0}{n_0}) + (\frac{n_j-n_0}{n_0})]$ is the fractional coordination number relative to the coordination number of the reference structure n_0, averaging between the coordination numbers n_i and n_j of atoms i and j. Using a function in a form as the screening function describe above, the coordination number can be calculated smoothly, $n_i = \sum_j(1 - S_{ij})$. Note that when i and j are nearest neighbour atoms, S_{ij} is close to 0 and n_i counts almost one neighbour. On the other hand, S_{ij} is close to one if i and j are not nearest neighbour atoms, so that n_i counts only a small fraction of a neighbour. By choosing the parameters for S_{ij} as $\beta_1=2.0$, $\beta_2=0.02895$, and $\beta_3=7.96284$, n_i are calculated to be 2.08469, 3.13698, 4.31114, 6.13943, 7.23413, 10.00215, and 11.93036 for the linear-chain, graphite, diamond, simple-cubic, β-tin, bcc, and fcc structures respectively. These values give a reasonable representation of the effective coordinations of these structures.

Besides the hopping parameters, the diagonal matrix elements are also dependent on the bonding environments. The expression for the diagonal matrix elements is

213

$$e_{\lambda,i} = e_{\lambda,0} + \sum_j \Delta e_\lambda(r_{ij}) \tag{5}$$

where $\Delta e_\lambda(r_{ij})$ takes the same expression as Eq. (1), λ denotes the possible types of orbitals (s, p, or d).

Finally, we express the repulsive energy term in a functional form as in the previous tight-binding model for carbon developed by Xu $et\ al.$ [7], that is

$$E_{rep} = \sum_i f(\sum_j \phi(r_{ij})), \tag{6}$$

where $\phi(r_{ij})$ is a pairwise potential between atoms i and j and f is a functional expressed as a 4th order polynomial with argument $x = \sum_j \phi(r_{ij})$, i.e., $f(x) = \sum_{n=0}^{n=4} c_n x^n$.

The parameters in the model are determined by first fitting to the first-principles density functional calculations results of electronic band structures and then the cohesive energy versus volume curves of crystalline structures with various coordinations. Additional checks have also been made to ensure that the model gives reasonable results for the elastic constants and phonon frequencies in the low energy structures. The details of the fitting procedure and parameters obtained from such fitting are described elsewhere [19, 20, 21].

RESULTS

Carbon, Silicon, and Germanium

Based on the model described in the previous section, we have generated environment tight-binding potentials for the semiconducting materials of C, Si, and Ge. The cohesive energies as a function of nearest neighbour distance for carbon, silicon, germanium in different crystalline structures obtained from the present tight-binding potentials are plotted in Fig.1 in comparison with the first-principles calculation results. The agreement between the present tight-binding calculation results and the LDA results is very good. These results demonstrate that the present environment-dependent tight-binding model has good transferability to describe not only the lower coordinated covalent structures, but also the higher coordinated metallic structures in this material. In contrast, two-center tight-binding models fail to describe covalent structures and metallic structures simultaneously. [7]

The present tight binding potentials also describe reasonably well the phonon frequencies and elastic constants of the materials. This can be seen from Tables I and II where the vibration frequencies of zone center and zone boundary modes and the elastic constants obtained by the present tight-binding potentials are compared with experimental data.

In addition to the properties of the bulk crystalline structures, we have performed extensive tight-binding molecular dynamics simulation to evaluate the accuracy and transferability of the new carbon and silicon potentials for describing more complex systems. Our studies show that these potentials also describe accurately the structures and energies of the diamond and silicon surfaces and silicon grain boundaries as well as those of the point defects in crystalline silicon. Details of these studies will be published in separate publications.
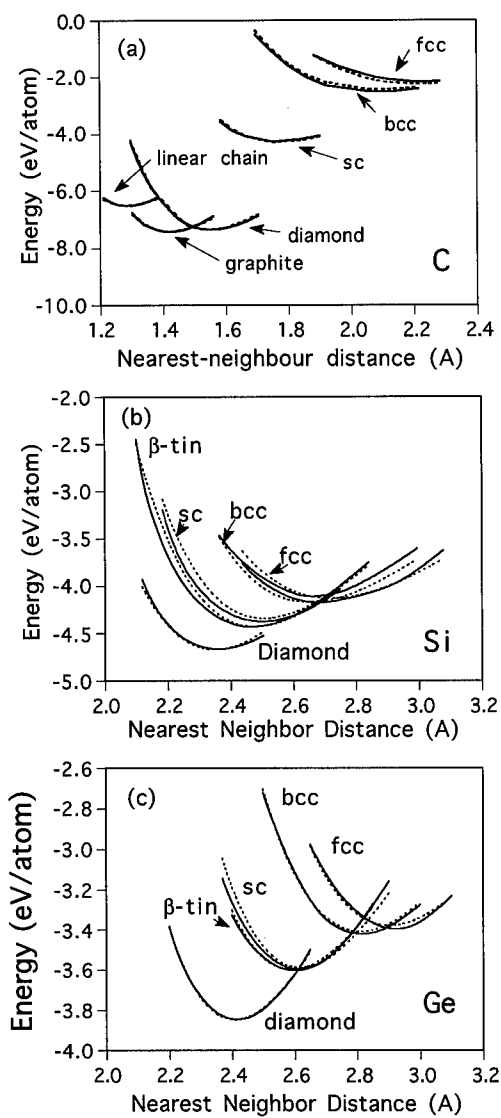
214

Figure 1: The cohesive energies as a function of nearest neighbour distance for carbon, silicon and germanium in different crystalline structures calculated using the present TB model are compared with the results from the first-principles LDA calculations. The solid curves are the TB results and the dashed curves are the LDA results.

Properties	Carbon		Silicon		Germanium	
	TB	Experiment	TB	Experiment	TB	Experiment
Lattice Constant (Å)	3.585	3.567	5.445	5.430	5.561	5.567
Bulk Modulus	4.19	4.42	0.902	0.978	0.774	0.772
$c_{11} - c_{12}$	9.25	9.51	0.993	1.012	0.703	0.801
c_{44}	5.55	5.76	0.716	0.796	0.602	0.666
$\nu_{LTO(\Gamma)}$	41.61	39.90	16.17	15.53	9.49	9.02
$\nu_{TA(X)}$	25.73	24.20	5.01	4.49	2.50	2.38
$\nu_{TO(X)}$	32.60	32.0	12.84	13.90	7.89	8.17
$\nu_{LA(X)}$	36.16	35.5	11.51	12.32	6.50	7.14

Table 1: Elastic constants and phonon frequencies of diamond, silicon, and germanium calculated from the present TB model are compared with and experimental results [22]. Elastic constants are in units of 10^{12}dyn/cm^2 and the phonon frequencies are in terahertz.

	Present Model	Experiment
$c_{11} - c_{12}$	8.94	8.80
E_{2g_2}	48.99	47.46
A_{2u}	26.07	26.04

Table 2: Elastic constants and phonon frequencies of graphite calculated from the present TB model are compared with experimental results [23]. Elastic constants are in units of 10^{12}dyn/cm^2 and the phonon frequencies are in terahertz.

<u>Aluminum, Niobium, and Molybdenum</u>

We have also applied the model to generating tight-binding potentials for simple metal system (Al) and bcc transition metal systems (Nb and Mo). Although the work is still in progress, preliminary results seem to be very encouraging.
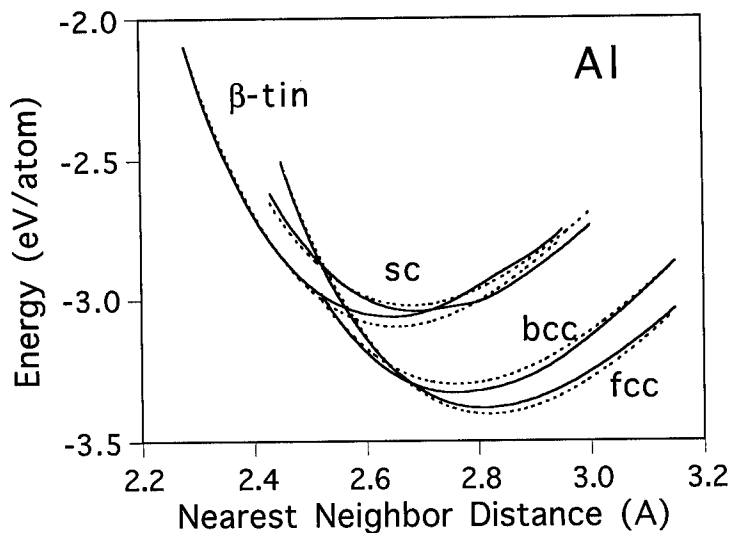
Figure 2: The cohesive energies as a function of nearest neighbour distance for aluminum in different crystalline structures calculated using the present TB model are compared with the results from the first-principles LDA calculations. The solid curves are the TB results and the dashed curves are the LDA results.

In Fig. 2, we show the energy versus nearest neighbor distance for aluminum in fcc, bcc, simple cubic, and β-tin structures. The results obtained from the present tight-binding calculations (solid line) fit well to the *ab initio* calculation results (dashed line). The phonon frequencies and elastic constants from the tight-binding calculation are also in good agreement with experimental data.

Our model can also describe the electronic band structures and energies of bcc transition metals with an orthogonal sp^3d^5 basis. Shown in Fig. 3 are the energies of Nb and Mo in different crystal structures as a function of lattice constant. Once again, the results from the present tight-binding model are found to be in very good agreement with the *ab initio* calculation results. More details about the tight-binding potential for molybdenum can be found in another report in this symposium.
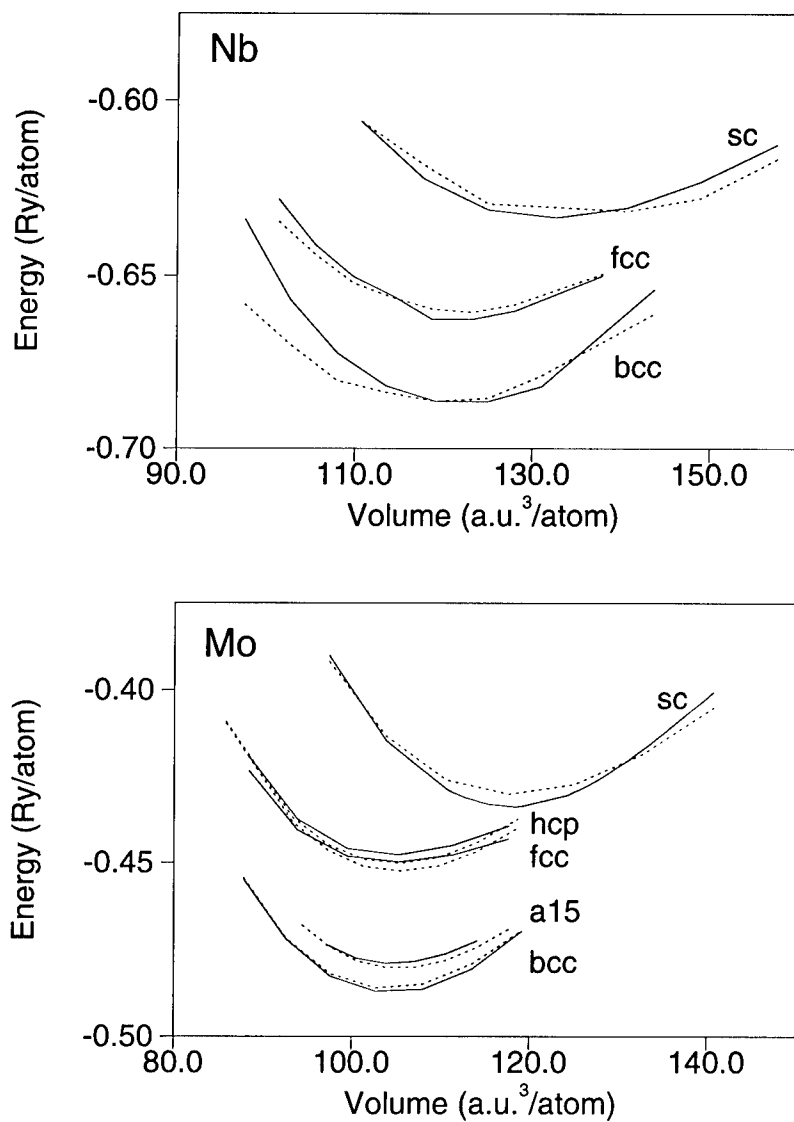
Figure 3: The total energies as a function of atomic volume for niobium and molybdenum in different crystalline structures calculated using the present TB model are compared with results from the first-principles LDA calculations. The solid curves are the TB results and the dashed curves are the LDA results.

CONCLUDING REMARKS

We outline in this paper a tight-binding model which goes beyond the traditional two-center approximation and allow the tight-biding parameters to change according to the bonding environment. We show that this model improves the transferability of the tight-binding potential to describe the structures and properties of metallic as well as covalent systems. For simplicity, our potentials are constructed using orthogonal minimal basis for C, Si, Ge, and Al and with additional five d-orbitals for the bcc transition metal elements. We argue that the effects of nonorthogonality, the effects of three-center integrals, and the effects of the elasticity of the basis can be, to some extent, taken into account by the environment-dependent scaling functions in our model. The environment-dependent scaling functions are introduced empirically in the present work. There is still room for improvement in the model if the form of these functions can be determined with the help of *ab initio* calculations. Another crucial issue in controlling the quality of the tight-binding parameters is the choice of fitting database. In the present work, our database involves mostly the properties of crystalline structures. It seems that the accuracy and transferability of the potentials would be further improved if an appropriate database which samples a variety of binding environments of the realistic materials is used to determine the parameters.

ACKNOWLEDGMENTS

REFERENCES

1. C. Z. Wang and K. M. Ho, Adv. Chem. Phys. vol. *XCIII*, 651 (1996).

2. X.-P. Li, R. W. Nunes, and D. Vanderbilt, Phys. Rev. B, **47**, 10891 (1993).

3. M. S. Daw, Phys. Rev. B, **47**, 10895 (1993).

4. F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993).

5. S. Y. Qiu, C. Z. Wang, K. M. Ho, and C. T. Chan J. Phys: Condens. Matter **6**, 9153 (1994).

6. L. Goodwin, A.J. Skinner, and D.G. Pettifor, Europhys. Lett. **9**, 701(1989).

7. C. H. Xu, C. Z. Wang, C. T. Chan, and K. M. Ho, J. Phys: Condensed Matter, **4**, 6047 (1992).

8. S. Sawada, Vacuum **41**, 612(1990).

9. M. Kohyama, J. Phys: Condens. Matter **3**, 2193 (1991).

10. I. Kwon, R. Biswas, C. Z. Wang, K. M. Ho, and C. M. Soukoulis, Phys. Rev. B **49**, 7242 (1994).

11. J. L. Mercer, Jr. and M. Y. Chou, Phys. Rev. B **47**, 9366 (1993).

12. J. L. Mercer, Jr. and M. Y. Chou, Phys. Rev. B **49**, 8506 (1994).

13. R. E. Cohen, M. J. Mehl, and D. A. Papaconstantopoulos, Phys. Rev. B **50**, 14694 (1994).

14. K. Stokbro, N. Chetty, K. W. Jacobsen, and J. K. Norskov, Phys. Rev. B **50**, 10727 (1994); K. Stokbro, Ph.D. thesis, Danish Technical University, Lyngby, De nmark, August 1994.

15. J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

16. Q. M. Li and R. Biswas, Phys. Rev B **50**, 18090 (1994).

17. O. K. Andersen and O. Jepsen, Phys. Rev. Lett. **53**, 2571 (1984).

18. K. Stokbro, K. W. Jacobsen, J. K. Norskov, D. Deaven, C. Z. Wang, and K. M. Ho, Surf. Sci. **360**, 221 (1996).

19. M. S. Tang, C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B**53**, 979 (1996).

20. B. C. Pan, C. Z. Wang, and K. M. Ho, *to be published.*

21. H. Haas, C. Z. Wang

22. *Semiconductors: Physics of Group IV Elements and III-V Compounds,* Landolt-Börnstein New Series III/17a, Eds. O. Madelung, M. Schulz, and H. Weiss, (Springer-Verlag, 1982); *Semiconductors: Intrinsic Properties of Group IV Elements and III-V, II-VI and I-VII Compounds,* Landolt-Börnstein New Series III/22a, Eds. O. Madelung, M. Schulz, (Springer-Verlag, 1987).

23. M. S. Dresselhaus and G. Dresselhaus, in *Light Scattering in Solids III,* ed. M. Cardona and G. Guntherodt (Springer, Berlin, 1982) p.8.

TIGHT-BINDING HAMILTONIANS FOR CARBON AND SILICON

D. A. PAPACONSTANTOPOULOS, M. J. MEHL, S. C. ERWIN, AND M. R.
PEDERSON
Complex Systems Theory Branch, Naval Research Laboratory, Washington DC
20375-5000

ABSTRACT

We demonstrate that our tight-binding method – which is based on fitting the energy bands and the total energy of first-principles calculations as a function of volume – can be easily extended to accurately describe carbon and silicon. We present equations of state that give the correct energy ordering between structures. We also show that quantities that were not fitted, such as elastic constants and the band structure of C_{60}, can be reliably obtained from our scheme.

INTRODUCTION

Over the past decade, many authors have developed tight-binding parametrizations for carbon [1, 2] and silicon [3, 4, 5, 6, 7, 8]. Because these parametrizations lead to small size Hamiltonians, they are particularly useful for the study of large unit cell systems such as the fullerenes [1, 9].

We have developed another tight-binding total energy method [10, 11] wherein the parameters are chosen to reproduce both the first-principles total energy and electronic band structure of face-centered and body-centered cubic crystals at several volumes. This method has been shown to work quite well in transition metals, where it correctly predicts the correct ground state and lower lying crystal states for all of the non-ferromagnetic transition metals. This includes the hcp metals and the α phase of manganese [12], even though these structures were not included in the fit. We also predict elastic constants, phonon frequencies, surface energies, and vacancy formation energies in agreement with experiment and first-principles results. The method has recently been extended to cover some of the sp bonded metals in column IIIA of the periodic table, aluminum, gallium, and indium [13]. For these elements the first-principles database had to be extended to include other than fcc and bcc crystals, but it did not include the body-centered tetragonal indium ground state or the complex αGa structure. Nevertheless, the method correctly predicts the ground state properties of all of these materials. The method also generates an electronic density of states for αGa which is very close to the results of first-principles calculations. [14]

Can this method be extended even further to the right of the periodic table, specifically to carbon and silicon? In this paper we develop a slightly modified version of the parametrization scheme developed in Ref. [11], applied to carbon and silicon. In the sections below we introduce the method, and discuss the properties predicted by our parameters.

METHOD

The tight-binding parametrization of Refs. [10, 11] consists of a prescription for specifying the behavior of the on-site parameters as a function of the local environment, and

Table I: The Slater-Koster tight-binding parameters for carbon, generated from the database described in the text.

On-Site Parameters (Eqs. (1) and (2))			
λ	1.59901905594		
Orbital	α (Ry)	β (Ry)	γ (Ry)	χ (Ry)
s	-0.102789972814	-1.62604640052	-178.884826119	4516.11342028
p	0.542619178314	2.73454062799	-67.139709883	438.52883145
Hopping Parameters (Eqs. (3))			
Orbital	a (Ry)	b (Ry/Bohr)	c (Ry/Bohr^2)	d (Bohr$^{-1/2}$)
$H_{ss\sigma}$	74.0837449667	-18.3225697598	-12.5253007169	1.41100521808
$H_{sp\sigma}$	-7.9172955767	3.6163510241	1.0416715714	1.16878908431
$H_{pp\sigma}$	-5.7016933899	1.0450894823	1.5062731505	1.13627440135
$H_{pp\pi}$	24.9104111573	-5.0603652530	-3.6844386855	1.36548919302
Overlap Parameters (Eqs. (4))			
Orbital	p (Bohr^{-1})	q (Bohr^{-2})	r (Bohr^{-3})	s (Bohr$^{-1/2}$)
$S_{ss\sigma}$	0.18525064246	1.56010486948	-0.308751658739	1.13700564649
$S_{sp\sigma}$	1.85250642463	-2.50183774417	0.178540723033	1.12900344616
$S_{pp\sigma}$	-1.29666913067	0.28270660019	-0.022234235553	0.76177690688
$S_{pp\pi}$	0.74092406925	-0.07310263856	0.016694077196	1.02148246334

a parametrization of the hopping and overlap matrix elements. We begin by discussing the behavior of the on-site terms, which are allowed to vary depending upon the local environment of each atom. This environment is determined by defining a pseudo-atomic density for each atom,

$$\rho_i = \sum_j e^{-\lambda^2|\mathbf{R}_j - \mathbf{R}_i|} f(|\mathbf{R}_j - \mathbf{R}_i|) , \tag{1}$$

where $\mathbf{R}_k$ is the position of the kth atom, the sum is over all neighbors of atom i, and where $f(R)$ is a cutoff function as defined in Ref. [11]. In this paper we choose the cutoff so that $f(R)$ vanishes when $R > 10.5$ atomic units for carbon, and when $R > 12.5$ atomic units for silicon. The on-site terms on each atom are given by a Birch-like equation

$$h_{i\ell} = \alpha_\ell + \beta_\ell \rho_i^{2/3} + \gamma_\ell \rho_i^{4/3} + \chi_\ell \rho_i^2 . \tag{2}$$

In both carbon and silicon we naturally consider only $\ell = s, p$, so, including the λ in equation (1), there are a total of nine parameters which determine the on-site terms on each atom.

The two-center Slater-Koster hopping terms for the Hamiltonian are simply polynomials times an exponential cutoff,

$$H_{\ell\ell'\mu}(R) = (a_{\ell\ell'\mu} + b_{\ell\ell'\mu}R + c_{\ell\ell'\mu}R^2) \exp(-d_{\ell\ell'\mu}^2 R) f(R) , \tag{3}$$

where f(R) is the same cutoff as above. The overlap parameters have been modified from our previous work[10, 11] to exhibit the proper behavior as the atoms get close together. Thus we write

$$S_{\ell\ell'\mu}(R) = (\delta_{\ell\ell'} + p_{\ell\ell'\mu}R + q_{\ell\ell'\mu}R^2 + r_{\ell\ell'\mu}R^3) \exp(-s_{\ell\ell'\mu}^2 R) f(R) , \tag{4}$$

Table II: The Slater-Koster tight-binding parameters for silicon, generated from the database described in the text.

On-Site Parameters (Eqs. (1) and (2))			
λ	1.10356625153		
Orbital	α (Ry)	β (Ry)	γ (Ry)	χ (Ry)
s	-0.053233461902	-0.907642743185	-8.83084913674	56.5661321469
p	0.357859715265	0.303647693101	7.09222903560	-77.4785508399

Hopping Parameters (Eqs. (3))			
Orbital	a (Ry)	b (Ry/Bohr)	c (Ry/Bohr^2)	d (Bohr$^{-1/2}$)
$H_{ss\sigma}$	219.560813651	-16.2132459618	-15.5048968097	1.26439940008
$H_{sp\sigma}$	10.127687621	-4.4036811240	0.2266767834	0.92267194054
$H_{pp\sigma}$	-22.959028107	1.7207707741	1.4191307713	1.03136916513
$H_{pp\pi}$	10.265449263	4.6718241428	-2.2161562721	1.11134828469

Overlap Parameters (Eqs. (4))			
Orbital	p (Bohr^{-1})	q (Bohr^{-2})	r (Bohr^{-3})	s (Bohr$^{-1/2}$)
$S_{ss\sigma}$	5.157587186	0.660009308	-0.0815441307	1.10814448800
$S_{sp\sigma}$	8.873646665	-16.240770475	5.1822969049	1.24065238343
$S_{pp\sigma}$	11.250489009	-1.1701322929	-1.0591485021	1.13762861032
$S_{pp\pi}$	-692.184231145	396.153248956	-13.8172106270	1.57248559510

where $\delta_{\ell\ell'}$ is the Kronecker delta function. In this paper we use only s and p orbitals, so we need only consider the Slater-Koster parameters $(\ell\ell'\mu) = (ss\sigma), (sp\sigma), (pp\sigma)$ and $(pp\pi)$.

The fitting procedure is described in [11]. Since we only consider s and p orbitals, we have the parameter λ from eq. (1); eight parameters $\alpha_\ell, \beta_\ell, \gamma_\ell$ and χ_ℓ from eq. (2); sixteen parameters $a_{\ell\ell'\mu}, b_{\ell\ell'\mu}, c_{\ell\ell'\mu}$ and $d_{\ell\ell'\mu}$ from eq. (3); and sixteen parameters $p_{\ell\ell'\mu}, q_{\ell\ell'\mu}, r_{\ell\ell'\mu}$ and $s_{\ell\ell'\mu}$ from eq. (4). This gives us a total of forty-one parameters. While this is a much larger number of parameters than used by some other methods [5, 7], this method is designed to map the electronic structure as well as the total energy information from a series of first-principles calculations onto the Slater-Koster parametrization, requiring a larger number of parameters for an accurate fit.

For each element, a database of first-principles density functional calculations is constructed. For carbon, this database contains the the one-electron eigenvalues as a function of k-point and volume, along with the total energy as a function of volume, for the diamond, graphite, and simple cubic lattice. In addition we include the eigenvalues and total energy of the C_2 dimer as a function of atomic separation. For silicon we fit only to the diamond, simple cubic, fcc, and bcc structures. In each case we shift the eigenvalues by a structure-dependent and volume dependent constant V_0, so that the total energy of a given structure S at a volume V is given by

$$E(\mathcal{S}, V) = \sum[\varepsilon_i + V_0(\mathcal{S}, V)] = \sum \varepsilon_i' \, , \qquad (5)$$

where the sums are over the occupied states. The tight-binding parameters are then chosen so as to reproduce both the total energies and the shifted eigenvalues ε_i' as closely as possible. A justification of this shift was recently proposed by McMahan and Klepeis.[8] The final tight-binding parameters of our scheme for carbon and silicon are listed in Tables I and II, respectively.
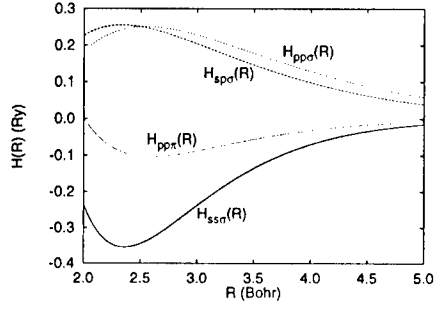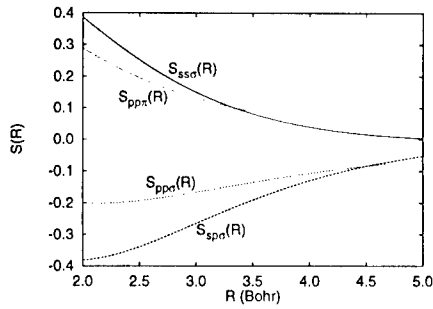
Figure 1: Two center Slater-Koster hopping matrix elements for carbon, based on the fitting procedure described in the text.

RESULTS

Carbon

We obtained our tight-binding parameters for carbon by fitting to a set of first-principles band structures and total energies, as outlined in Ref. [11]. Results for the bulk diamond, graphite, and the simple cubic structures were obtained using the full-potential Linearized Augmented Plane Wave (LAPW) method,[15, 16] and the Wigner form of the Local Density Approximation. We also found it desirable to include results for the C_2 molecule in our database. These calculations used the all-electron density functional based NRLMOL codes [17, 18, 19]. In the dimer there is a crossing between the π_u state and the $3\sigma_g$ state near the experimental equilibrium distance.[19] We dealt with this problem by fitting only in the regions where the ordering of the states is unambiguous. Thus we fit to the fully

Figure 2: Two center Slater-Koster overlap matrix elements for carbon, based on the fitting procedure described in the text.

π-bonded molecule at short bondlengths, and to the partially $(\pi_u, 3\sigma_g)$-bonded state at large bondlengths. The final tight-binding parameters for carbon are shown in Table I.

In Fig. 1 we show the two-center Slater-Koster hopping matrix elements $ss\sigma$, $sp\sigma$, $pp\sigma$ and $pp\pi$ derived from our parameters in Table I as a function of the interatomic distance R. It should be noted that these parameters decrease smoothly to zero for $R \geq 10.5$ Bohr and also follow the correct sign convention, that is, $ss\sigma$ and $pp\pi$ are negative, while $sp\sigma$ and $pp\sigma$ are positive.[20] In Fig. 2 we show the distance dependence of the corresponding overlap matrix elements. In the interaction region ($R > 2$ Bohr) each matrix element has the opposite sign from its respective hopping matrix element, as it should.[20] However, note that the overlap matrix element $pp\sigma$ must approach 1 as the distance approaches zero, as required by (4) and on physical grounds. Finally, note that the overlap matrix elements also approach zero at large distances.

In Fig. 3 we show our TB results for the total energy of carbon as a function of the nearest-neighbor distance in several structures. It is clear that our tight-binding Hamiltonian works very well; it shows the cubic diamond and graphitic plane structures with the lowest energy, and the hexagonal diamond structure nearly degenerate with cubic diamond. It does not, however, predict the correct ground state of graphite. When we vary the volume and c/a ratio of the graphite structure to find the minimum energy configuration, we find that the structure stabilizes into a set of infinitely separated graphite planes with an in-plane lattice constant $a = 2.472$Å. This occurs because our current fit does not adequately account for the interaction between the graphite planes. For reference, the experimental equilibrium lattice constant for graphite is $a = 2.464$Å.[21]

From various first-principles calculations we estimate that the energy of the C_{60} molecule is about 0.4eV/atom above the energy of the diamond lattice. Using our TB parameters we find this energy to by 0.38eV/atom. It should be emphasized that we did not fit to the
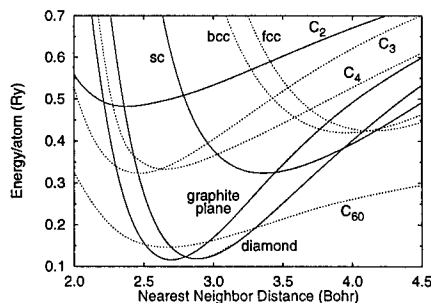
Figure 3: Energy of carbon predicted by our tight-binding parameters. Several different structures are shown, with the relative energy per atom shown as a function of the nearest neighbor separation for that structure. All other internal and external parameters were allowed to relax to minimize the total energy. The structures marked by solid lines were used in the fit, while the structures marked by dotted lines are predictions of the method. The unlabeled curve for the hexagonal diamond structure is nearly degenerate with the cubic diamond curve.

TB   LDA

Energy (eV)
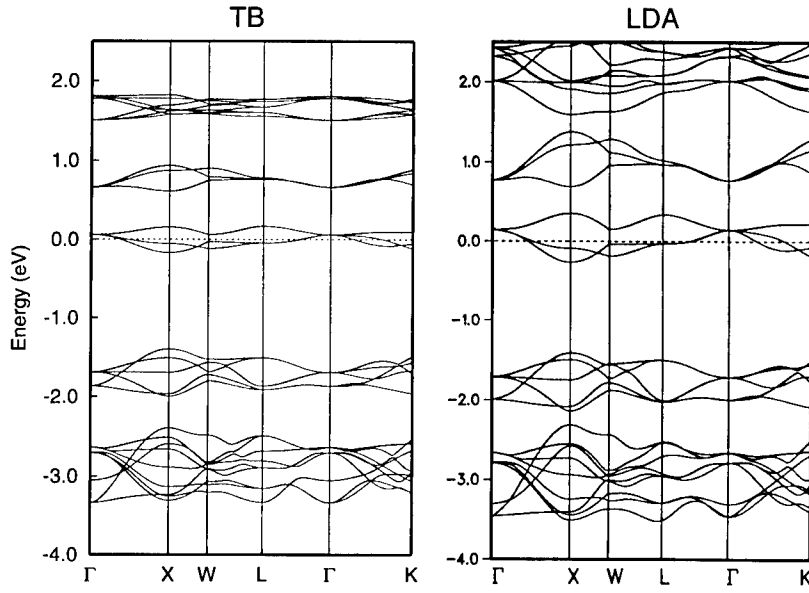
Figure 4: The left graph shows TB band structure of fcc C_{60}, compared to a the first-principles band structure of K_3C_{60} determined in Ref. [22]. The Fermi level in both pictures is set to correspond to a system with three electrons in the conduction band, as in K_3C_{60}. and is then set to zero.

C_{60} molecule and that this result is an output of our TB method. We further investigated the behavior of C_{60} by calculating the energy bands of an fcc crystal of C_{60} molecules, with an eye to studying the superconducting phase transition in K_3C_{60}.[22, 23] In Fig. 4 we show the band structure of an fcc crystal of C_{60} molecules calculated with our tight-binding parameters, and compare it to a first-principles calculation for K_3C_{60}. Our TB valence and conduction bands compare quite favorably with the LDA results, although the TB bands are slightly narrower. We note that the first band above the Fermi level is significantly narrower than the LDA, and the next band is somewhat lower in energy than the LDA. However, in superconductivity the important band is the threefold band crossing the Fermi level. These bands are well reproduced in TB which will allow us to carry out further studies of the electronic density of states in this system for different orientational phases of the C_{60} molecules.

We determined the elastic constants of diamond at the experimental lattice constant by imposing a finite strain on the crystal, calculating the total energy, and obtaining the corresponding elastic constant from the curvature of the energy-strain relation at zero strain.[25, 26] In calculating the bulk modulus $(C_{11}+2C_{12})/3$ and tetragonal shear $C_{11}-C_{12}$

Table III: The elastic constants of diamond structure carbon and silicon at the indicated room temperature lattice constants, using the TB parameters of Tables I and II. For comparison, we also show the results of the full-potential LAPW calculations described in the text, and the experimental values.[24] Since the LAPW calculations only give an upper bound for C_{44}, we also give the upper bound found using our TB parameters.

		Carbon a = 3.567Å			Silicon a = 5.430Å		
		TB	LAPW	Exp.	TB	LAPW	Exp.
B		377	420	442	108	91	98
$C_{11} - C_{12}$		988	926	951	106	92	102
C_{11}		1036	1037	1076	179	152	166
C_{12}		48	111	125	73	60	64
C_{44}	Upper Bound	601	555		135	101	
C_{44}	Relaxed	601		576	95		80

we can impose strains so that the diamond basis remains fixed by symmetry. In calculating the shear constant C_{44}, however, we must allow the carbon atoms to move as we strain the lattice. For comparison, we also calculated the elastic moduli using the LAPW method with the Wigner parametrization of the local density approximation. In this case we did not allow the carbon atoms to move, so we can only calculate an upper bound on C_{44}. When we calculated this upper bound from the TB parameters, however, we found that there is very little relaxation, as shown in Table III. Our results are shown in Table III. The TB parameters underestimate the bulk modulus by about 15% compared to experiment (10% compared to the LAPW result) and overestimate $C_{11} - C_{12}$ by 4% (7%). These combine to produce the very small result for C_{12}, which is 60% lower than experiment (57% compared to LAPW). The shear modulus C_{44} is 4% above the experimental value (8% above LAPW).
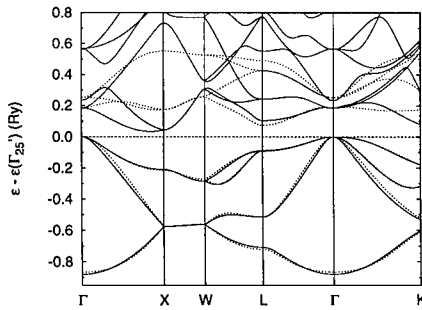
Figure 5: The electronic band structure of silicon at the experimental lattice constant, 5.430Å. The solid lines represent the results of the first-principles LAPW method. The dashed lines show the results obtained from the TB parameters in Table II.

227
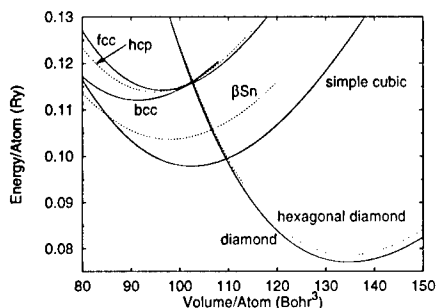
Figure 6: The TB equation of state for several structures silicon, using the parameters given in Table II. The structures marked by solid lines were used in the fit, while the structures marked by dotted lines are predictions of the method.

Silicon

We performed similar first-principles calculations for silicon using the LAPW method and the Hedin-Lundqvist prescription for exchange and correlation. We fitted our TB Hamiltonian parameters to four structures (diamond, simple cubic, fcc, and bcc) using a non-orthogonal sp basis. In Fig. 5 we show a comparison of TB and LAPW results for theband structure. It is evident that the valence bands fit very well. However, the conduction bands are poorly fit, and in TB the smallest gap is at the symmetry point L instead of on the Δ line near X. We have done further work using an spd representation which dramatically improves the conduction band. These results will be reported elsewhere. The present fit is still useful because by restricting the TB to s and p orbitals we have a more convenient model for use in molecular dynamics simulations. Results of similar quality were recently reported by Cohen et al.[6] using the same approach but with some differences in the form of the on-site terms.

In Fig. 6 we plot the equation of state for various phases of silicon. In this plot the hexagonal diamond, βSn, and hcp phases are predictions of the method. This plot should be compared to the first-principles results of Yin and Cohen.[27] We see that we have the correct ordering of all structures except the βSn state, which is too high in energy.

In Table III we compare the calculated TB and first-principles elastic constants of silicon to experiment. Again, the first-principles value of C_{44} is only an upper bound. We see that in this case the elastic constants of silicon are nearly as accurate as the first-principles results. Note that there is a considerable change in C_{44} when we relax the internal parameters.

SUMMARY

We have proposed an extension of our parameterized tight-binding method – which was originally formulated for highly coordinated metals – to the more open structures of carbon and silicon. We are able to accurately describe the relative energies of several

different structural phases, the elastic constants of the ground state phase, and the occupied electronic band structure relative to first-principles methods. Future work will focus on the electronic structure and total energies for geometries far outside the fitting database, including silicon surface reconstructions and the C_{60} orientational phase diagram.

ACKNOWLEDGMENTS

REFERENCES

1. M. Menon and K. R. Subbaswamy, Phys. Rev. Lett. **67**, 3487 (1991).

2. M. S. Tang, C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B **53**, 979 (1996).

3. Th. Frauenheim, F. Weich, Th. Köhler, and S. Uhlmann, Phys. Rev. B **52**, 11492 (1995).

4. T. J. Lenosky, J. D. Kress, I. Kwon, A. F. Voter, B. Edwards, D. F. Richards, S. Yang, and J. B. Adams, Phys. Rev. B **55**, 1528 (1997).

5. M. Menon and K. R. Subbaswamy, Phys. Rev. B **55**, 9231 (1997).

6. R. E. Cohen, L. Stixrude, and E. Wasserman, Phys. Rev. B **56**, 8575 (1997).

7. N. Bernstein and E. Kaxiras, Phys. Rev. B **56**, 10488 (1997).

8. A. K. McMahan and J. E. Klepeis, Phys. Rev. B **56**, 12250 (1997).

9. D. Porezag, G. Jungnickel, T. Frauenheim, G. Seifert, A. Ayuela A, M. R. Pederson, Applied Physics A – Materials Science and Processing **64**, 321 (1997).

10. R. E. Cohen, M. J. Mehl, and D. A. Papaconstantopoulos, Phys. Rev. B **50**, 14694 (1994).

11. M. J. Mehl and D. A. Papaconstantopoulos, Phys. Rev. B **54**, 4519 (1996).

12. M. J. Mehl and D. A. Papaconstantopoulos, Europhysics Letters **31**, 537 (1996).

13. S. Yang, M. J. Mehl, and D. A. Papaconstantopoulos, Phys. Rev. B, to appear (1997).

14. All of the tight-binding parameters based on Refs. [11] and [13], as well as the parameters used in this paper, are available on the World Wide Web at http://cst-www.nrl.navy.mil/bind/.

15. O. K. Andersen, *Phys. Rev. B* **12**, 3060 (1975).

16. S. H. Wei and H. Krakauer, *Phys. Rev. Lett.* **55**, 1200 (1985).

17. M. R. Pederson and K. A. Jackson, Phys. Rev. B **41**, 7453 (1990)

18. M. R. Pederson and K. A. Jackson, Phys. Rev. B **42**, 3276 (1990).

19. M. R. Pederson and K. A. Jackson, Phys. Rev. B **43**, 7312 (1991).

20. W. A. Harrison, *Electronic Structure and the Properties of Solids*, (Dover, New York, 1989).

21. P. Villars and L.D. Calvert, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases*, 2nd Edition (ASM International, Materials Park, Ohio, 1991), p. 1854.

22. S. C. Erwin and W. E. Pickett, Science **254**, 842 (1991).

23. S. Satpathy, V. P. Antropov, O.K. Andersen, O. Jepsen, O. Gunnarsson, and A. I. Liechtenstein, Phys. Rev. B **46**, 1773 (1992).

24. H. L. Anderson, Ed., *A Physicist's Desk Reference, The Second Edition of Physics Vade Mecum*, (American Institute of Physics, New York, 1989).

25. M. J. Mehl, Phys. Rev. B **47**, 2493 (1993).

26. M. J. Mehl, B. A. Klein, and D. A. Papaconstantopoulos, in *Intermetallic Compounds: Principles and Applications*, J. H. Westbrook and R. L. Fleischer, eds. (Wiley, London, 1994). Vol. I, Chapter 9.

27. M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 5668 (1982).

ELECTRONIC STRUCTURE AND TRANSPORT IN NON PERIODIC SYSTEMS: NEW O(N) METHODS

D.Mayou*, P. E. A. Turchi**, S. Roche*** and J. P. Julien+**
** LEPES-CNRS, B.P.166, F-38042 Grenoble Cedex 9, France; ** LLNL (L-268), PO Box 808, Livermore CA 94551; *** Dept of Applied Physics, University of Tokyo, Bunkyo-ku, Tokyo 113 Japan; +** Université Française du Pacifique, Papeete, Tahiti.

ABSTRACT

The mathematical theory of orthogonal polynomials and continued fractions provides efficient tools, via the recursion and related methods, for calculating diagonal elements of Green's function of tight-binding Hamiltonians. We present two recent generalizations of this formalism. The first one allows the calculation of conductivity and other linear response coefficients. The second one provides a new approach to the solution of mean-field theories of alloys. In particular it is shown that the self-consistent CPA equations can be easily solved, through a real-space calculation, for multi-component alloys based on periodic or non periodic lattices.

INTRODUCTION

It has long been recognized that structural order affects a majority of the physical properties exhibited by most alloys as functions of temperature, pressure and concentration. Models based on tight-binding Hamiltonians have been developed to calculate electronic structure and predict for example energetic or transport properties. However, in many cases the periodicity of an underlying lattice is assumed. In the context of quasicrystalline or amorphous alloys, and in particular recently synthetized bulk amorphous alloys, there is a need to develop efficient electronic structure methods, entirely solvable in real space, and thus applicable to systems with reduced or no symmetry.

The aim of this paper is to review recent progress in that field. The developments presented here are all based on the mathematical theory of orthogonal polynomials and continued fraction expansions, which are related to the recursion method [1]. In a first part we recall some basic results concerning the recursion method. Then we show how one can derive a useful expression for the projector of states of energy E for a tight-binding Hamiltonian H. This expression allows the calculation of various quantities like dc or ac conductivity, or forces acting on an atom in a molecular dynamics simulation. In the last section we show on a simple example that the recursion method provides a new real-space solution to mean-field theories of alloys. Application to the Coherent Potential Approximation (CPA) is given which proves the interest of this approach in particular in solving the inhomogeneous CPA problem in the case of multi-component alloys based on non periodic lattices.

THE RECURSION BASIS AND SOME USEFUL RELATIONS

The recursion basis in tight-binding models

We consider a system described by a tight-binding Hamiltonian H. To a given normalized state $|\Psi_1 >$ we can always associate a recursion basis which is constructed as follows. If we consider $H|\Psi_1 >$ we can decompose it in a component parallel to $|\Psi_1 >$ and a component orthogonal to $|\Psi_1 >$. Thus we write

$$H|\Psi_1 >= a_1|\Psi_1 > +b_1|\Psi_2 >$$

where $< \Psi_1|\Psi_2 >= 0$ and $< \Psi_2|\Psi_2 >= 1$. If the coefficient b_1 is chosen real and positive then $|\Psi_2 >$ is defined in a unique way. Consider then $H|\Psi_2 >$, this vector can be decomposed in a component parallel to the space spanned by $|\Psi_1 >$, $|\Psi_2 >$ and a component orthogonal to this space. We get

$$H|\Psi_2 >= a_2|\Psi_2 > +b_1'|\Psi_1 > +b_2|\Psi_3 >$$

with $< \Psi_1|\Psi_3 >= 0$ $< \Psi_2|\Psi_3 >= 0$, and $< \Psi_3|\Psi_3 >= 1$. Furthermore since H is a hermitian operator we deduce that $b_1' = b_1$. If the coefficient b_2 is chosen real and positive then $|\Psi_3 >$ is defined in a unique way.

The process can be repeated and leads to the construction of a set of states $|\Psi_n >$ which are orthonormal and satisfy :

$$H|\Psi_n >= a_n|\Psi_n > +b_{n-1}|\Psi_{n-1} > +b_n|\Psi_{n+1} >$$

In the basis of $|\Psi_n >$ the Hamiltonian H is tridiagonal. This means that the Hamiltonian is that of a semi-linear chain as shown in Figure 1.
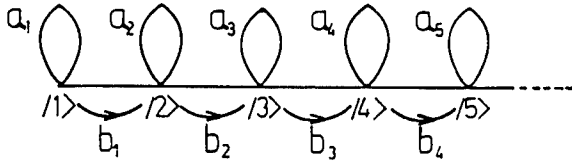
Figure 1: Representation of the semi-infinite linear chain of the recursion basis. The coefficients a_i are the on-site energies of sites i and the coefficients b_i are the hopping terms between nearest-neighbor sites i and $i+1$.

An important property of the states $|\Psi_n >$ is that they spread progressively from the initial state. For example if $|\Psi_1 >$ is located on an atomic orbital then the state $|\Psi_n >$ can have non-zero components only on orbitals that are reached by $n-1$ applications of the Hamiltonian H. This is illustrated in Figure 2 for the simple case of a single s-band model applied to a two dimensional lattice with nearest neighbor hopping.
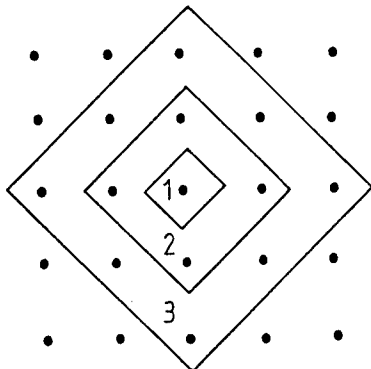
Figure 2: Extension of the wavefunctions of the recursion basis for a square lattice (s-band with nearest neighbor hopping). The initial orbital is one of the s-orbitals of the lattice.

The interest of the recursion basis is that if we want to calculate vectors like $f(H)|\Psi_1 >$ where $f(H)$ is any operator that is a function of the Hamiltonian H then this vector can be decomposed in the basis of the states $|\Psi_n >$:

$$f(H)|\Psi_1 >= \sum_n |\Psi_n >< \Psi_n|f(H)|\Psi_{=1} >$$

Hence we have to calculate $< \Psi_n|f(H)|\Psi_1 >$ for a one dimensional Hamiltonian H with known matrix elements a_n and b_n.

When applied to the resolvent $f(H) = (z - H)^{-1}$ one can show in particular that

$$< \Psi_1|\frac{1}{z-H}|\Psi_1 >= \cfrac{1}{z-a_1-\cfrac{b_1^2}{z-a_2-\cfrac{b_2^2}{z-a_3-....}}}$$

This continued fraction expansion of the diagonal element of the resolvent provides an efficient way of reconstructing the density of states [1].

The recursion basis in functional space

Consider a function $n(E)$ positive definite such that $\int dE n(E) = 1$, and which is non-zero only in a finite interval. From $n(E)$ one can define a scalar product in the space of functions of E. Namely

$$< f|g >= \int dE n(E) f^*(E) g(E)$$

One can also define an operator that we note $\mathcal{H}$. The action of this linear operator is to multiply a function $f(E)$ by E. This operator is hermitian. One can show also that :

$$n(E) =< \Psi_1|\delta(E - \mathcal{H})|\Psi_1 >$$

where $|\Psi_1 >= 1$, that is, the function associated with this state is a constant equal to 1. This vector is normalized due to the normalization condition of the function $n(E)$. In the vectorial space spanned by the functions $f(E)$ one can apply the recursion procedure described above starting from $|\Psi_1 >= 1$.

The recursion procedure applied to this initial vector will generate a set of functions which are polynomials of E. The vector $|\Psi_n >$ is a polynomial of degree $n - 1$ that we note $P_n(E)$. It can be shown that these polynomials are a basis for the functions of the energy E in the energy interval where $n(E)$ is non-zero. Thus for any, sufficiently regular, function $f(E)$ we have :

$$|f >= \sum_n |\Psi_n >< \Psi_n|f >$$

that is

$$f(E) = \sum_n P_n(E) < \Psi_n|f >$$

The polynomials $P_n(E)$ obey a three-term recurrence relation analogous to that of the recursion vector in the case of tight-binding models. In particular one can show that the recursion vectors $|\Psi_n >$ for the tight-binding case is equal to $P_n(H)|\Psi_1 >$, and the $P_n(E)$ are the orthogonal polynomials associated with the density of states $n(E)$ and the state $|\Psi_1 >$.

CALCULATION OF LINEAR RESPONSE COEFFICIENTS

Linear response theory provides a suitable formalism for determining important characteristics of a system, such as conductivity, Hall effect etc.... The force applied to an atom in a

molecular dynamics simulation is also a linear response (the derivative of the energy with respect to the displacement). Many of these quantities can be expressed in terms of Green's function or equivalently in term of the operator $\delta(E - H)$. We show here that one can get, through the theory of orthogonal polynomials, a useful expression for the operator $\delta(E - H)$ [2]. This expression can serve as a starting point for evaluating the linear response of a system.

Consider the function of energy $f(E)$ given by

$$f(E) = \frac{<\alpha|\delta(E-H)|\beta>}{n(E)}$$

where the electronic density of states $n(E)$ is non-zero on the spectrum of H, and $|\alpha>$ and $|\beta>$ are two wavefunctions. From the previous section, we know that $f(E)$ can be developed on the basis of the orthogonal polynomials associated with the density $n(E)$. If we apply the decomposition formula we get :

$$f(E) = \sum_n P_n(E) < \Psi_n|f> = \sum_n P_n(E) < \alpha|P_n(H)|\beta>$$

Since this equality holds for any pair of states $|\alpha>$ and $|\beta>$, we can write formally :

$$\delta(E - H) = n(E)\sum_n P_n(E)P_n(H)$$

This expression has been used to calculate the time evolution of a wavepacket and the dc conductivity in quasiperiodic systems [3]. One uses the Chebyshev polynomials of the first kind :

$$
\begin{aligned}
e^{-\frac{iHt}{\hbar}}|j> &= \sum_n h_n(\int dEn(E)Q_n(E)e^{-\frac{iEt}{\hbar}}) \ Q_n(H)|j> \\
&= \sum_n h_n i^n J_n(\frac{-2b_\infty t}{\hbar}) \ e^{\frac{-ia_\infty t}{\hbar}} Q_n(\mathcal{H})|j>
\end{aligned}
$$

where $h_0 = 1$ and $h_n = 1/2$ for $n > 0$. As usually done with orthogonal polynomials the $Q_n(\mathcal{H})|j>$ are evaluated via the recurrence property [1]. Amplitudes of $e^{-\frac{iHt}{\hbar}}|j>$ in this basis have a rather simple form connected to Bessel functions. The consequent interest of this approach is that one obtains well conditioned asymptotic behaviors for these coefficients which converge very quickly as $n \to \infty$, given that $\lim_{n\to\infty} J_n(z) \sim \frac{1}{\sqrt{2\pi n}}(\frac{ez}{2n})^n$. It is this development that makes the calculation fast and accurate.

This method is very efficient for studying problems related to quantum diffusion or conductivity. Its application allowed us to show the specificity of electronic transport in quasicrystals and to exhibit the different type of plausible conduction mechanisms. The method is particularly efficient for calculating the dc conductivity even in systems with long mean-free paths of the order of several tens of interatomic distances. The above development has also been used in a different way to calculate the ac and dc conductivity in disordered systems [2].

It would be interesting to compare our method to the one used by Bose [4] for computing the dc conductivity. We note that in his method Bose calculates the conductivity for each Fermi energy whereas in our method we obtain the conductivity at all energies in one run.

Finally we mention that the efficient tight-binding molecular dynamic method proposed by Goedecker and Colombo [5] can be viewed as a direct application of the spectral decomposition of $\delta(E - H)$ on the basis of the orthogonal polynomials. Indeed for a component (say along the x-direction) of the force acting on an atom :

$$f_x = \sum_{\alpha_{occupied}} < \alpha|\frac{dH}{dx}|\alpha> = \text{Tr}[\frac{dH}{dx} \int dE\delta(E - H)]$$

where the integral is over energies less than the Fermi energy. Using the decomposition of $\delta(E - H)$ we get :

$$f_x = \sum_n C_n \text{Tr}[\tfrac{dH}{dx} Q_n(H)]$$

and

$$C_n = \int dE n(E) Q_n(E)$$

where again the integral is over energies less than the Fermi energy. If we take for $Q_n(E)$ the Chebyshev polynomials we recover the method of Goedecker and Colombo [5].

NEW APPROACH TO MEAN FIELD THEORIES OF ALLOYS

A simple case. Principle of the approach

Let us consider a s-band model applied to the case of an alloy based on a square lattice with nearest neighbor hopping only. In order to illustrate the approach we consider an approximation in which the infinite square lattice is replaced by a cluster which is a 3x3 square [6] .This is a simple approximation to describe the environment of the central atom of the cluster. We take into account the fact that the eight surface atoms of the cluster are coupled to the rest of the lattice in the following way. To the "i" atom we attribute a self-energy $\sigma_i(z)$ which is $\sigma_i(z) = \sigma(z)$ for an atom of the surface which is coupled to one atom of the rest of the lattice, and $\sigma_i(z) = 2\sigma(z)$ for an atom of the surface which is coupled to two atoms of the rest of the lattice (i.e., atoms at the corners of the 3x3 square).

In the spirit of a self-consistent model we can relate $\sigma(z)$ simply to the self-energy $\Delta(z)$ of the atom at the center of the cluster. We choose the relation

$$\Delta(z) = 4\sigma(z)$$

since $\Delta(z)$ describes the coupling of the central atom to four neighbors.

The resulting set of equations is defined as follows. The Hamiltonian $H(\sigma(z))$ of the 3x3 cluster depends on the self-energy $\sigma(z)$. One has with $|1>$ meaning the central orbital

$$\tfrac{1}{z-\Delta(z)} = <1|\tfrac{1}{z-H(\sigma(z))}|1>$$

This equation (I) gives an expression for $\Delta(z)$ as function of $\sigma(z)$ due to the form of the effective medium (here a cluster with surface atoms carrying a self-energy $\sigma(z)$). We call this equation (I) the effective medium equation. The other equation (equation (II)) $\Delta(z) = 4\sigma(z)$ stems from the self-consistency condition that we have chosen. We call equation (II) the self-consistency equation.

Usually one would solve these equations by iteration for a given value of z. One starts, for example, from a value of $\sigma(z)$ then one calculates the diagonal element of the resolvent and deduces $\Delta(z)$ from equation (I). From equation (II) one deduces a new value for $\sigma(z)$. This new value is used as a starting point for the next iteration. One reiterates the procedure until convergence is reached within some given accuracy .

Within this new approach, we proceed in a completely different way. Since $\sigma(z)$ represents the coupling to the rest of the medium it must be a Herglotz function which can be represented by a continued fraction,

$$\sigma(z) = \cfrac{b_0{}^2}{z-a_1-\cfrac{b_1{}^2}{z-a_2=-\cfrac{b_2{}^2}{z-a_3-....}}}$$

Thus it is strictly equivalent to give a self-energy $N\sigma(z)$ ($N = 1$ or $N = 2$) to a surface atom or to couple it with a semi-linear chain with coefficients $\sqrt{N}b_0$, a_1, b_1, a_2, b_2, a_3 ...This newly defined effective medium is equivalent to the system represented in Figure 3.

Hence the principle for solving the set of equations in a self-consistent way follows. Because of equation (II) (the self-consistency equation) we get a simple relation between the continued fraction coefficients of $\Delta(z)$ and the continued fraction coefficients of $\sigma(z)$. They are identical except that b_0 for $\sigma(z)$ is replaced by $2b_0$ for $\Delta(z)$. Thus one makes a recursion calculation starting from the central orbital $|1>$. The essential point stems from the extension of the successive recursion vectors as represented in Figure 3.

Indeed, if we consider $|2>$ we see that it is entirely in the 3x3 square and thus can be calculated without any knowledge of the chains coefficients. From this calculation one deduces the first two coefficients of the continued fraction coefficients of $\Delta(z)$: that is, one calculates b_0 and a_1. Once these two coefficients are known we see from the extension of $|3>$ that the knowledge of the Hamiltonian represented in Figure 3 is sufficient for defining this vector. From this calculation we deduce b_1 and a_2 and so on.

Application to the CPA
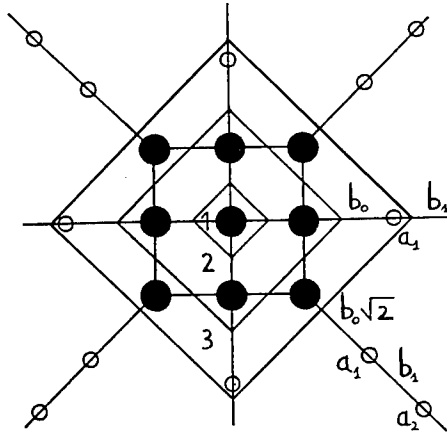
Figure 3: Representation of the effective Hamiltonian for the cluster embedded in an effective medium. The coupling to the semi-linear chains is equivalent to the effect of the self-energy attibuted to each atom of the surface of the cluster. The extension of the successive recursion vectors $|1>$, $|2>$, $|3>$ starting from the central orbital $|1>$ is represented.

In the case of the CPA the principle is the same as above but the details are more intricate [7]. In the CPA it is well known that the chemical disorder is simulated by attibuting a self-energy $\sigma(z)$ to each atom of the medium surrounding an A or a B atom (for a binary alloy). As illustrated in Figure 4, one can attach semi-linear chains to the atoms of the medium. These semi-linear chains will simulate exactly the self-energy $\sigma(z)$ provided that the coefficients of the Hamiltonian of the chains are the coefficients of the continued fraction expansion of $\sigma(z)$. Note that this expansion exists since $\sigma(z)$ is Herglotz.
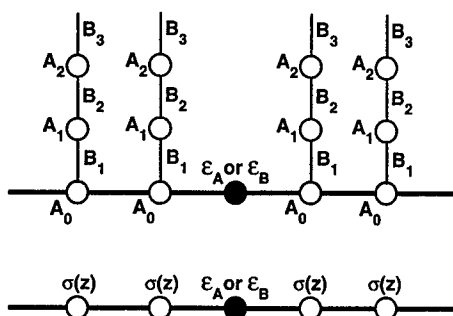
Figure 4: Equivalent representations of the effective Hamiltonian describing chemical disorder, within the CPA, here for a binary alloy based on an infinite linear chain (thick solid lines).

One can also define the self-energy $\Delta(z)$ of an orbital centered on an A or a B atom due to the coupling with the effective medium. This self-energy $\Delta(z)$ is thus a function of $\sigma(z)$. This corresponds to the effective medium equation (I). The equation corresponding to the self-consistency condition (equation (II)) results from the condition that the T-matrix of an A or a B atom embedded in the effective medium is zero on average. It can be shown that this condition gives a relation between the continued fraction coefficients of $\sigma(z)$ and $\Delta(z)$. This relation is more complex than in the above example and is given in reference [8]. The main point is that, as in the above example, it allows one to perform a recursion starting from an orbital centered on an A or a B atom in the effective medium. One calculates progressively the recursion vectors and the coefficients of the semi-linear chains as illustrated in Figure 4. At each step of the recursion procedure one knows enough coefficients to calculate the next recursion vector, and so on. The extension of the first recursion vectors starting from an orbital is represented in Figure 5.

This procedure to solve the CPA equations appears to be quick and furthermore it does not require a periodic lattice. It can be easily generalized to several orbitals per atom and several inequivalent sites as in the case of a surface or a topologically disordered lattice. Also, since one
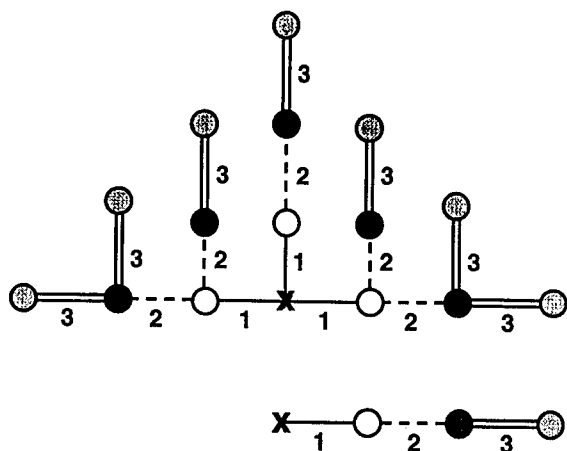
Figure 5: Equivalent representations of $G_{eff}(z)$ by a lattice, here an infinite linear chain with each site "dressed" by a semi-linear chain representing $\sigma(z)$ (top part), and by a semi-linear chain in the recursion basis (bottom part). The extension of the recursion vector at each step, here $1 - 3$, is represented by solid, dashed, and double solid lines, respectively.

237

has now an effective, energy independent Hamiltonian, it is possible to use formalisms developed in particular for tight-binding Hamiltonians. This allowed us to calculate effective pair interactions between atoms with the orbital peeling method of Burke [9]. This also allows us also to calculate for example transport properties in chemically and topologically disordered systems with the real-space techniques described above [10].

The finite electronic lifetime, that results from chemical disorder, leads also to another interesting aspect of the method. Indeed, let us consider a wavevector in the effective, energy independent Hamiltonian (see Figure 4), and suppose that this wavevector has initially non-zero components only on sites of the real system (not on sites of the chains). It is easy to show that during its evolution in time this state will spread on the chains and its total weight on the sites of the real system will decrease exponentially with a characteristic time τ equal to the electronic finite lifetime induced by disorder. Consequently if the state is initially located in a small region of space (an atomic orbital for example) it will have time to travel at most on a distance of a few mean free-path l, and then will spread into the chains attached to the atoms of this region. The same will happen for the recursion vectors calculated from an atomic orbital. In other words a recursion calculation requires the knowledge of the local environment of an orbital only on a distance of a few times the mean-free path l. This is illustrated in Figure 6 where, in the case of a chemically disordered alloy, we show that results obtained by simulating exactly only the first three shells of neighbors of an atom are close to the exact results.

CONCLUSION

To conclude we stress that the new methods presented here are all real-space methods; consequently they are applicable to any system, periodic or not, and perfectly suited for implementation on massively parallel computers. They provide a consistent set of techniques for calculating electronic structure and transport properties of non periodic systems.
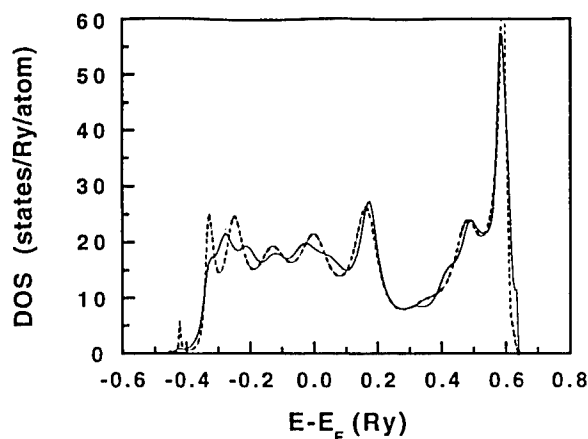
Figure 6: Illustration of the fact that the recursion procedure requires a description of the local environment on a distance of the order of the electronic mean-free path. Density of states of bcc-based ZrRh disordered alloy; (i) Obtained with 17 levels of continued fraction on both the real crystal and the semi-linear chain (solid line). (Note that this calculation is essentially exact); (ii) Obtained with three levels of continued fraction on the real crystal and ten levels along the semi-linear chains.

REFERENCES

1. R. Haydock, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic Press, New York, 1980), Vol **35**.

2. D. Mayou, Europhys. Lett.**6**, 549 (1988); D. Mayou and S. Khanna, J. Phys. I (France) **5**, 1199 (1995).

3. S. Roche and D. Mayou, Phys. Rev. Lett. **79**, 2518 (1997).

4. S. K. Bose, in these proceedings.

5. S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994).

6. D. Mayou, Thesis Université Joseph Fourier Grenoble (1987), unpublished.

7. J. P. Julien and D. Mayou, J. Phys. I (France) **3**, 1861 (1993).

8. P. E. A. Turchi, D. Mayou, and J. P. Julien, Phys. Rev. B **56**, 1726 (1997).

9. N. R. Burke, Surf. Sci. **58**, 349 (1976).

10. R. Girardi, P. E. A. Turchi, and D. Mayou, to be published.

LET THERE BE LIGHT IN TIGHT BINDING

P. VOGL[*], M. GRAF[*], A. GÖRLING[**]
[*]Walter Schottky Institute, Techn. Univ. Munich, D-85748 Garching, Germany,
vogl@wsi.tu-muenchen.de
[**]Institute of Theoretical Chemistry, Techn. Univ. Munich, D-85748 Garching,
Germany

ABSTRACT

Empirical **tight-binding** theory is generalized to incorporate time-dependent
electromagnetic fields in a systematic and gauge-invariant manner that does not
introduce any extra adjustable parameters. It is shown that this approach successfully
predicts a wide range of solid state properties that have not been accessible within
the tight binding method so far. We present applications such as optical constants,
luminescence in heterostructures, properties in ultra-high magnetic fields and **lattice
dynamics** in polar materials.

INTRODUCTION

Empirical tight-binding theory has become an indispensable tool for obtaining
both physical insight as well as semi-quantitative predictions for a variety of elec-
tronic or structural properties of solids. Since the pioneering work of Harrison [1],
empirical tight-binding method has been widely used in many different areas such
as bulk band structures [2-4], surfaces [5], transition metals [6], cohesive properties
[7,8], lattice dynamics [9], molecular dynamics [10]. So far, the vast majority of appli-
cations of tight-binding theory has focused on the spectrum of the electronic Hamil-
tonian in the absence of any external fields. Relatively few attempts have been made
to incorporate electromagnetic fields into the theory in a systematic fashion [11-14].

In this paper, we first discuss a sum rule that provides a link between tight-
binding theory and the linear response formalism [15]. We then show that time-
dependent electromagnetic fields $\Phi(r,t)$, $A(r,t)$ can be incorporated in empirical tight
binding theory in a way that is gauge invariant, charge conserving, and, most impor-
tantly, does not introduce any extra free parameters [15]. This scheme significantly
widens the scope of tight binding theory and the range of problems it can be success-
fully applied to. We illustrate this novel scheme by applying it to optical **dielectric
functions**, **optical properties** of heterostructure devices, **magnetic band structures**,
dynamical effective charges, and phonon spectra in polar materials.

THEORY

Notation

In the empirical tight binding scheme [1], the Hamiltonian matrix is represented
in terms of an orthogonal basis set of atomic-like orbitals $|\alpha,R\rangle$ that can be character-
ized by a site index R and a symmetry-related index α that specifies the angular-
momentum and spin quantum numbers of the atomic orbitals on that site. The

intra- and inter-atomic Hamiltonian matrix elements are denoted by, respectively,

$$\varepsilon_{\alpha,R} = \; <\alpha,R \mid H \mid \alpha,R> \; , \tag{1}$$

$$t_{\alpha',\alpha}(R' - R) = \; <\alpha',R' \mid H \mid \alpha,R> \; . \tag{2}$$

In all applications of this paper, we will invoke the two-center approximation and restrict the range of interactions to nearest neighbors. Tight binding models of this kind have the major advantage of (i) requiring a minimal number of adjustable parameters that can be determined largely by physical considerations and (ii) being easily transferable in the sense that parameters for compounds AB and AC suffice to calculate, e.g., heterostructures AB:AC or alloys AB_xC_{1-x}. Spin-orbit interaction can be taken into account in tight binding by adding intra-atomic Hamiltonian matrix elements [16] of the form $t_{\alpha'\alpha}(0) \propto <\alpha',R \mid L \cdot S \mid \alpha,R>$ but we will not discuss relativistic effects here.

In a periodic solid, one can construct Bloch *basis* functions $\mid\alpha,k>$ from a linear combination of all atomic orbitals of type α in the crystal. The Bloch *eigenfunctions* of the Hamiltonian H are denoted by $\mid nk>$, where the index n labels the energy bands. The expansion coefficients of the eigenstates in terms of the Bloch basis functions $\mid\alpha,k>$ define the coefficient vector $C(nk)$,

$$H_{\alpha\alpha'}(k) = \; <\alpha',k \mid H \mid \alpha,k> \tag{3}$$

$$\mid nk> = \sum_{\alpha} C_{\alpha}(nk) \mid \alpha,k> \; . \tag{4}$$

Effective mass sum rule

In standard $k \cdot p$ theory, one expresses the matrix elements $<n'k \mid H \mid n \; k>$ in terms of matrix elements made up entirely of eigenstates and eigenvalues at some reference wave vector k^*, using the completeness of the Kohn-Luttinger basis. Even though the Wannier-type basis used in tight binding theory is incomplete, one can still derive an exact result for $<n'k \mid H \mid n \; k>$ that resembles the well-known $k \cdot p$ result [15]. To this end, we expand Eq. (3) in powers of $(k - k^*)$. This gives, in matrix notation,

$$H(k) = \; H(k^*) + \nabla_{k^*}H(k^*) \cdot (k-k^*) + \frac{1}{2}(k-k^*)\cdot\nabla_{k^*}\nabla_{k^*}H(k^*)\cdot(k-k^*) + O(k-k^*)^3 \; . \tag{5}$$

By defining an *effective* momentum operator p and an *effective* kinetic energy operator T (m_0 is the free electron mass),

$$p_{nn'}(k^*) = \frac{m_0}{\hbar} C^{\dagger}(nk^*) \; \nabla_{k^*} H(k^*) \; C(n'k^*) \; , \tag{6}$$

$$T_{nn'}(k^*) = \frac{m_0}{\hbar^2} C^{\dagger}(nk^*) \; \nabla_{k^*}\nabla_{k^*} H(k^*) \; C(n'k^*) \; , \tag{7}$$

it is easy to see that the matrix elements of $H(k)$ in the Bloch basis $|n\ k^*\rangle$ can be written as [15]

$$H_{nn'}(k) = E_n(k^*)\,\delta_{n,n'} + \frac{\hbar}{m_0}\,p_{nn'}(k^*)\cdot(k-k^*) + \frac{\hbar^2}{2m_0}\,(k-k^*)\cdot T_{nn'}(k^*)\cdot(k-k^*)$$

$$+ (k-k^*)\cdot\frac{\hbar^2}{m_0^2}\sum_{m\neq n,n'}\frac{p_{nm}(k^*)\,p_{mn'}(k^*)}{E_n(k^*)-E_m(k^*)}\cdot(k-k^*)\ . \tag{8}$$

This result has a similar form as the standard expression for $H_{nn'}(k)$ in $k\cdot p$-theory. Recently, it has been rederived in a more general framework [17]. Note that the intra-band term $T_{nn'}$ equals the unity operator $\delta_{nn'}$ in the Kohn-Luttinger basis. For a nondegenerate band edge $E_n(k^*)$, Eq. (8) yields the *identity*

$$\frac{1}{\hbar^2}\frac{\partial^2 E_n(k^*)}{\partial k_i^* \partial k_j^*} = \frac{1}{m_0}\,T_{nn}(k^*)_{ij} + \frac{1}{m_0^2}\sum_{m\neq n,n'}\frac{p_{nm}^i(k^*)\,p_{mn}^j(k^*) + p_{nm}^j(k^*)\,p_{mn}^i(k^*)}{E_n(k^*)-E_m(k^*)}\ . \tag{9}$$

This result not only gives the effective mass but actually constitutes a finite-basis analogue of the optical f-sum rule [15]. For degenerate bands, Eq. (8) yields expressions for the Luttinger parameters that differ from the standard ones through the operator $T_{nn'}$. For simple tight binding models, these band edge parameters and effective masses can be evaluated analytically [18].

Electromagnetic fields

We turn to the incorporation of electromagnetic potentials into the tight-binding approach. It is certainly plausible to assume that an electrostatic potential $\Phi(R,t)$ produces a rigid shift of all on-site valence orbital energies at position R and does not – or not significantly – affect the off-site matrix elements. Indeed, this approximation has been widely used [1,11,19]. It is less obvious, however, how to incorporate a vector potential in a localized basis approach and how to ensure gauge invariance.

Given a general Hamiltonian $H(r, p)$ that is a function of the position and momentum operator, minimal coupling to a vector potential $A(r,t)$ can be achieved by the Peierls-transformation

$$\exp[-\frac{ie}{\hbar c}\int^r A(s,t)\cdot ds]\ H(r,p)\ \exp[\frac{ie}{\hbar c}\int^r A(s,t)\cdot ds]\ = H(r,p+\frac{e}{c}A(r,t))\ . \tag{10}$$

It turns out that a discrete variant of this expression – together with the approximation for $\Phi(R,t)$ that we just mentioned – yields a gauge invariant and charge conserving [15] tight binding Hamiltonian matrix. It reads

$$\varepsilon_{\alpha,R} = \varepsilon_{\alpha,R}^0 - e\,\Phi(R,t)\ , \tag{11}$$

$$t_{\alpha',\alpha}(R', R) = t_{\alpha',\alpha}^0(R' - R)\exp\left\{-\frac{ie}{2\hbar c}(R'-R)\cdot[\,A(R',t)+A(R,t)\,]\right\}\ , \tag{12}$$

where the field-free matrix elements are ε^0 and t^0. The electrostatic potential only enters the on-site matrix elements, whereas the vector potential enters only the off-site elements. We note that Eqs. (11), (12) neglect intra-atomic excitations that become relevant for very high excitation energies in molecules or solids and in isolated atoms.

Within one-band or k·p-models, external fields are sometimes incorporated by the Luttinger-Kohn approximation [20]. In this approach, the wave vector k in the Hamitonian matrix $H_{nn'}(k)$ is replaced by an operator $k \rightarrow -i\nabla + e\,A(r,t)/\hbar\,c$. This approximation breaks down whenever the external fields induce significant band mixing. Most of the examples that we present in the following subsections belong to this category.

APPLICATIONS AND RESULTS

Magnetic band structure

As a first demonstration for Eqs. (11), (12) to yield intuitively correct results, we study the spectrum of a two-dimensional square lattice (lattice constant a) with 1 single s-state per site in a magnetic field perpendicular to the two-dimensional plane. In this case, Eqs. (11), (12) yield the so-called Harper equation [21] that shows pronounced commensurability effects between the periodic potential and the cyclotron motion [15]. Whenever the flux $\Phi = Ba^2$ through the unit cell is a rational multiple of the flux quantum $\Phi_0 = hc/e$, the Hamiltonian becomes periodic and its spectrum consists of energy bands of finite width. Figure 1 shows this spectrum for $\Phi/\Phi_0 = p/q$ with $q \leq 30$.

In fact, any crystal Hamiltonian in a magnetic field becomes periodic whenever the ratio of magnetic flux through the unit and Φ_0 is a rational number [22]. However, the periodicity cell is extremely large for realistic fields. The corresponding spectrum has been termed magnetic band structure [22]. Except in the one-band case, it can neither be obtained within the Kohn-Luttinger approach nor is it accessible via perturbation theory. Therefore, only few attempts have been made so far to calculate the magnetic band structure of real materials [23]. For a magnetic field of ≈ 630 T along the [001] direction, the magnetic supercell of bulk GaAs contains 160 primitive cells perpendicular to the magnetic field. Consequently, the resulting magnetic
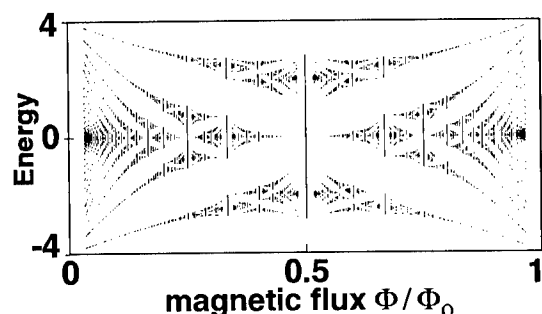

Figure 1: Energy bands of two-dimensional square lattice as a function of magnetic flux ("Hofstadter butterfly" [21]), as calculated from the Hamiltonian matrix elements Eqs. (11), (12).
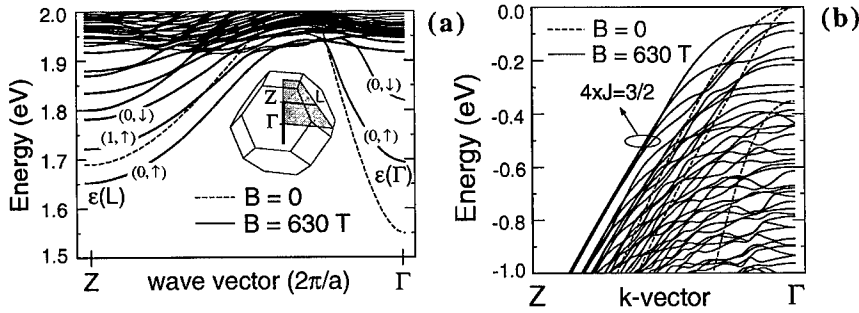
Figure 2: Calculated magnetic band structure of bulk GaAs in ultrahigh magnetic fields B. The inset shows the magnetic Brillouin zone (thick line) within the fcc-zone. (a) Lowest conduction bands from the Γ-point to Z; the lowest bands at Z originate from zero-field L-states. (b) Top valence band states.

Brillouin zone reduces practically to a line when plotted on the same scale as the Brillouin zone of the bulk lattice (see Fig. 2) and leads to a highly degenerate folded band structure along the [001]-direction.

We have calculated the magnetic band structure of bulk GaAs within the approach outlined above, i. e. Eqs. (11), (12). Figure 2a shows the calculated magnetic energy band structure of GaAs for the lowest conduction bands at a very high magnetic field, the top of the valence band (at B=0) representing the zero of energy [15]. The minimum of the conduction band at the Z point in the magnetic Brillouin zone originates in the L point of the underlying bulk zone. Since the effective mass at Γ is much smaller than at L, the magnetic field-induced zero point energy $\hbar\omega_c/2$ becomes so large at B=630 T that it raises the conduction band minimum at the Γ point *above* the L-minimum, leading to an indirect band gap material. Consequently, these calculations predict that GaAs becomes indirect at a magnetic field around 500 Tesla. It should be mentioned that such ultrahigh magnetic field pulses are actually attainable nowadays [in the high magnetic field laboratory of the University of Tokyo, see [24]].

Figure 2b depicts the same magnetic energy band structure of GaAs for the top valence band states. In this and the previous picture, we have employed a relativistic sp^3s^* tight binding model [25]. For energies below 0.3 eV of the valence band top, the constant energy surfaces are no longer spherical. This strong nonparabolicity effect, together with band mixing effects, causes the cyclotron energy to change non-monotonically with wave vector and leads to *negative* effective masses for some hole bands. The magnetic field induced spin-orbit splitting becomes so large at these high fields that the J=3/2 and J=1/2 band states no longer interact with one another. This leads to a fourfold degeneracy of the highest lying hole bands that is also indicated in the figure.

Transverse dielectric function

Once we know the change of the Hamiltonian matrix upon application of a vector potential, as given in Eqs. (11), (12), we can calculate the transverse dielectric tensor $\varepsilon_{ij}(\omega)$ as a function of frequency ω, invoking the standard procedures of linear

response theory. The indices i, j denote the cartesian components. In [15] it has been shown that Eqs. (11), (12) lead to the following expression,

$$\text{Re } \varepsilon_{ij}(\omega) = 1 - \frac{4\pi e^2}{\omega^2 \Omega} \sum_{nk} f_n(\mathbf{k}) \left(\frac{1}{m_n(\mathbf{k})}\right)_{ij} + \frac{4\pi e^2}{m_0^2 \Omega} \mathcal{P} \sum_{nmk} \frac{[f_n(\mathbf{k}) - f_m(\mathbf{k})] \; p_{nm}^i(\mathbf{k}) p_{mn}^j(\mathbf{k})}{\omega_{nm}^2(\mathbf{k}) \; \hbar(\omega_{mn}(\mathbf{k}) - \omega)} \; , \quad (13)$$

$$\text{Im } \varepsilon_{ij}(\omega) = \frac{4e^2 \pi^2}{\omega^2 m_0^2 \Omega} \sum_{nmk} [f_n(\mathbf{k}) - f_m(\mathbf{k})] \; p_{nm}^i(\mathbf{k}) \; p_{mn}^j(\mathbf{k}) \; \delta \, (\hbar\omega - \hbar\omega_{mn}(\mathbf{k}) \,) \; . \quad (14)$$

Here, $f_n(\mathbf{k})$ is the Fermi distribution function of the Bloch eigenstates $|nk\rangle$. The momentum operator $\mathbf{p}$ has been defined in Eq. (6). $\mathcal{P}$ denotes the principal value, we have set $\hbar\omega_{mn}(\mathbf{k})=E_m(\mathbf{k})-E_n(\mathbf{k})$, Ω is the crystal volume, and m_0 the free electron mass. The effective mass tensor $[1/m_n(\mathbf{k})]_{ij}$ can be deduced from Eq. (9). Interestingly, this expression for the transverse dielectric function is *formally identical* to the standard random phase expression [26], provided we interpret the operator $\mathbf{p}(\mathbf{k})$ as momentum operator. Note that the second term on the right hand side of Eq. (13) vanishes for an insulator with full valence bands. As we have pointed out before, however, this analogy does not hold in general due to the appearance of the kinetic energy operator $\mathbf{T}$ in Eq. (7) that differs from unity in any finite basis. In fact, the effective mass in Eqs. (9) and (13) is numerically dominated by the $\mathbf{T}$-matrix element for most energy bands in GaAs.

In Fig. 3, we show the calculated real and imaginary part of the dielectric function for intrinsic GaAs. We have used a set of $sp^3d^5s^*$ nearest neighbor tight-binding parameters that have been very carefully determined recently [27] and give excellent band gaps and effective masses. As one can see, agreement between theory and experiment [28] is very good, apart from a small shift in the peak positions. Fig. 3 also shows the joined density of states that equals – up to an arbitrary normalization factor – the expression Eq. (14) with $\mathbf{p}$ set equal to unity. This comparison shows that the momentum matrix elements determine the relative weight of the two peaks of ε_2 at 3 and 5 eV and the high energy tail of ε_2, all in excellent accord with experiment.
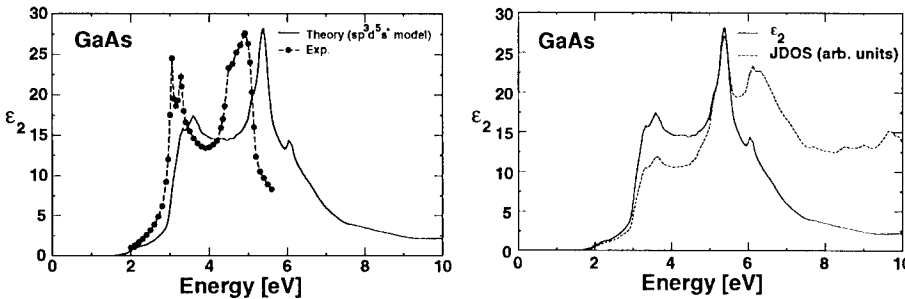

Figure 3: Left: Calculated imaginary part of dielectric function (full line) with tight binding parameters from Ref. [27]. Experimental data (dotted line) are from Ref. [28]. Right: Full line is same as on left side, the dashed line shows the joined density of states in arbitrary units.

Optical properties and charge self-consistency in heterostructures

We now turn to electronic and optical properties of semiconductor nanostructures. Specifically, we consider a lattice matched heterostructure on GaAs, consisting of an undoped GaAs buffer, a 26.5 nm wide strained layer $In_{0.15}Ga_{0.85}As$ quantum well, and a 38 nm wide n-doped $Al_{0.2}Ga_{0.8}As$ cap layer ($n=10^{18}$ cm^{-3}) the first 2 nm of which are undoped. This layer sequence is typical for a high electron mobility transistor (HEMT). The incorporation of electromagnetic fields into the tight binding approach as outlined above allows one to calculate the self-consistent charge distribution in this structure as a function of an applied (gate) voltage, the confined electronic states within the quantum well, as well as the luminescence spectrum [29].

Computationally, it is adequate in this case to take open boundary conditions well outside the $In_{0.15}Ga_{0.85}As$ channel since there is no current flow in growth direction. By introducing layer Bloch functions $|E\ k_{/\!/}\rangle$ with in-plane wave vector $k_{/\!/}$, the secular equation $H\ |E\ k_{/\!/}\rangle = E\ |E\ k_{/\!/}\rangle$ becomes effectively one-dimensional along the growth axis z. Here, $H = H_0 + V_H$, where H_0 is the tight-binding Hamiltonian of the intrinsic structure and $V_H(z)$ is the self-consistent Hartree potential that enters only in the on-site matrix elements of H, according to Eq. (11). The influence of electronic charge rearrangement can be included by solving Poisson's equation $d(\varepsilon(z)\ dV_H/dz)/dz = -4\pi\rho(z)$, where the charge density is

$$\rho(z) = \frac{e}{(2\pi)^2} \int dk_{/\!/} \sum_n |\langle nk_{/\!/} | z\rangle|^2 f(E_n) \qquad (15)$$

self-consistently with the secular equation for H. The index n labels all energy levels for given $k_{/\!/}$. In Fig. 4, we show the resulting band profiles and quantized subband states for zero bias and for a gate voltage of -0.6 V across the 200 monolayer heterostructure. Strain effects in the InGaAs region which are taken into account by properly scaling the matrix elements of H [25] split the heavy hole (HH) and light hole (LH) states. Note that the quantum well becomes almost flat for a bias of -0.6 V.

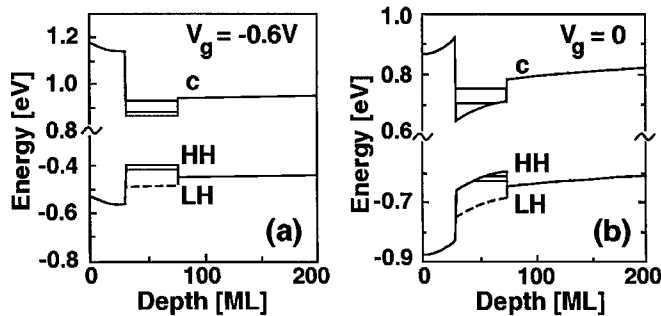Since Bloch symmetry is lost in z-direction, the expression Eq. (13), (14) for the

Figure 4: (a) Self-consistent band edge profile (thick lines for HH, dashed lines for LH), and bound states (thin lines) of pseudomorphic $Al_{0.2}Ga_{0.8}As$-$In_{0.15}Ga_{0.85}As$-GaAs HEMT structure at zero bias gate voltage V_g. The layer width is measured in monolayers [ML]. (b) Calculated photoluminescence spectra as a function of V_g.

247

transverse dielectric function is no longer adequate but has been properly generalized in [29]. Figure 5 shows the calculated luminescence emission rate as a function of applied bias. The change from 2 peaks to 4 peaks is a consequence of the breakdown of optical selection rules due to band mixing that is induced by the band bending within the quantum well for small bias.

Lattice dynamics and self-consistent tight-binding

In order to study cohesive properties, lattice dynamics, and longitudinal response functions in empirical tight binding theory, one needs to establish a relation between the net charges in the crystal and the potential these charges generate [7,11,25,30] In the present paper, we have employed a self-consistent sp^3 model that was developed in [7,31]. We first briefly summarize this model and then apply it to lattice dynamics.

To simplify notation, we lump together the symmetry and site index into a single index $I = (\alpha, R)$. A quite general Hartree ansatz for the on-site energies is [7]

$$\varepsilon_I = w_I - \sum_{I'} (Z_{I'} - Q_{I'})\, U(R_I - R_{I'}) + f_I \ , \tag{16}$$

where Z_{Ie} is the core charge of the ion at site R_I and w_I are the orbital energies of the neutral atoms. The electronic orbital occupancy on site R_I is given by

$$Q_I = \frac{1}{N} \sum_{nk\,I} f_n(k)\, |<I\,|nk>|^2 \tag{17}$$

so that $Q_I = Z_I$ for neutral atoms within the solid. The potential U reflects the on-site Coulomb repulsion for $R_I = R_{I'}$ and tends to a Coulomb potential for long distances and has also been given in [7]. Finally, the term $f_I = -(S H + H S)_{I,I}/2$ is a non-orthogonality correction term that contains orbital overlap factors $S_{I,I'}$. Again, we
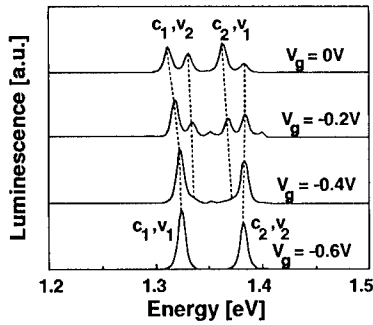

Figure 5: Photoluminescence of HEMT structure in Fig. 4 as a function of bias voltage V_g. The numbers label the two electron (c) and hole (v) bound states. The dashed lines mark corresponding transitions.
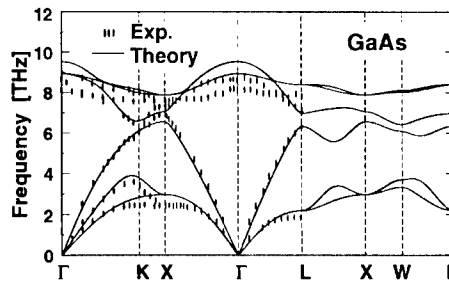
Figure 6: Calculated phonon spectra across the Brillouin zone of bulk GaAs, compared to experiment [32].

248

have used the expressions in [7], where $S_{I,I'} \propto d_{I,I'}^{-3}$ and $d_{I,I'}$ is the distance between neighboring atoms I and I'. Since the orbital energies in Eq. (16) depend, via Eq. (17), implicitly on the eigenstates $|nk\rangle$, they must be determined self-consistently. It has been shown [7,31] that Eq. (16) leads to the following expression for the cohesive energy of the crystal,

$$E_{coh} = E_{cov} + E_{ovl} + E_{transf}$$
(18)

$$E_{cov} = \sum_{nk} f_{nk} E_{nk} - \sum_{I} Q_I \epsilon_I \ ,$$

$$E_{ovl} = \sum_{I} Q_I f_I \ ,$$

$$E_{transf} = -\sum_{I} (Z_I - Q_I) w_I + \frac{1}{2} \sum_{I\,I'} (Z_I - Q_I)(Z_{I'} - Q_{I'}) \, U(R_I - R_{I'}) \ .$$

The three contributions to the cohesive energy have a simple physical interpretation. The first term, called covalent energy, is the lowering of the electronic energy due to the formation of bonds. The second contribution takes into account the increase of the total energy of the solid upon compression and has been termed overlap energy. The third term, called charge transfer contribution, contains all electrostatic terms and vanishes if there are no charged ions in the crystal.

This model allows one to study the response of the crystal to any static potential, in particular the one that is induced by ionic displacements of the form

$$\mathbf{u}_I = \mathbf{u}_{q\tau} \exp(i\mathbf{q} \cdot \mathbf{R}_I) \ ,$$
(19)

where the index τ denotes the sublattice associated with site I. Employing linear response theory [33], the force constants

$$\Phi_{ij}(\tau,\tau' \mid q) = \frac{1}{N} \, \partial^2 E_{coh} / \partial u_{q\tau}^{*i} \, \partial u_{q\tau'}^{*j}$$
(20)

can be calculated explicitly but the expressions are too lengthy to be given here. In Fig. 6, we show some numerical results for GaAs. Comparison with experiment reveals that the present theory fits the data amazingly well, considering the fact that no parameter in the tight-binding model of [7] was adjusted.

An elucidating side result of the calculation of the dynamical matrix is the transverse effective charge tensor $\mathbf{Z}^*$. It is defined as the macroscopic dipole moment $\mathbf{P}$ induced by homogeneous sublattice displacements $\mathbf{u}_\tau$ for zero macroscopic electric field [33],

$$\lim_{q \to 0} \mathbf{q} \cdot \mathbf{P}(q) \Big|_{E=0} = \frac{e}{\Omega_0} \sum_{\tau,ij} q_i \, Z_{ij}^*(\tau) \, u_\tau^j \ .$$
(21)

In the present tight binding approach, we obtain the following analytical expression,

$$Z^*_{ij}(\tau) = (Z_\tau - Q_\tau) \, \delta_{ij} - \frac{\hbar}{m_0} \sum_{nmk} \frac{f_{nk} - f_{mk}}{(E_{nk} - E_{mk})^2} \; p^i_{nm}(k) < mk \mid \frac{\partial H}{\partial u^j_\tau} \mid nk > \; , \qquad (22)$$

that closely resembles a corresponding general expression that has been derived within pseudopotential theory [34]. The so-called acoustic sum rule,

$$\sum_\tau Z^*_{ij}(\tau) = 0 \; . \qquad (23)$$

is obeyed individually by the first term and the second term on the right hand side of Eq. (22). The first term vanishes due to charge conservation in the crystal and the second due to the fact that the Hamiltonian is invariant under a rigid shift of the crystal. Numerically, we find $Z^* = 1.53$ in GaAs which underestimates the experimental value of 2.07 somewhat.

We note that dynamical effective charges can alternatively be calculated using the Berry phase approach which has recently also been applied to tight binding [35]. It does not give analytical expressions for Z^* but yields numerically very similar results [35].

ACKNOWLEDGMENTS

REFERENCES

1.   W. A. Harrison, Electronic Structure and the Properties of Solids (Freeman, San Francisco, 1980).

2.   P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Chem. Solids **44**, 365 (1983).

3.   D. A. Papaconstantopoulos, Handbook of the Band Structure of Elemental Solids (Plenum, New York, 1986).

4.   C. Priester, G. Allan, and M. Lannoo, Phys. Rev. B **37**, 8519 (1988).

5.   G. Lengel, R. Wilkins, G. Brown, M. Weimer, J. Gryko, and R. E. Allen, Phys. Rev. Lett. **72**, 836 (1994).

6.   W. A. Harrison and G. K. Straub, Phys. Rev. B **36**, 2695 (1987).

7.   J. A. Majewski and P. Vogl, Phys. Rev. B **35**, 9666 (1987).

8.   D.G. Pettifor, Solid State Physics **40**, 43 (1987).

9.   A. Mazur and J. Pollmann, Phys. Rev. B **39**, 5261 (1989).

10. O. F. Sankey and R. E. Allen, ibid. **33**, 7164 (1986); M. Menon and R. E. Allen, Phys. Rev. B **33**, 7099 (1986).

11. W. A. Harrison and J. E. Klepeis, Phys. Rev. B **37**, 864 (1988).

12. L. C. Lew Yan Voon, L. R. Ram-Mohan, Phys. Rev. B **47**, 15500 (1993).

13. H. M. Polatoglou, G. Theodorou, and C. Tserbak, Phys. Rev. B **49**, 8132 (1994).

14. J. C. Durán, F. Flores, C. Tejedor, and A. Munoz, Phys. Rev. B **36**, 5920 (1987).

15. M. Graf and P. Vogl, Phys. Rev. B **51**, 4940 (1995); M. Graf, PhD thesis, Techn. Univ. Munich, 1996.

16. D. J. Chadi, Phys. Rev. B **16**, 790 (1977).

17. T. B. Boykin, Phys. Rev. B **52**, 16317 (1995).

18. T. B. Boykin, G. Klimeck, R. C. Bowen, and R. Lake, Phys. Rev. B **56**, 4102 (1997).

19. A. Di Carlo and P. Vogl, Phys. Rev. B **50**, 8358 (1994).

20. J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).

21. D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).

22. E. Brown, Solid State Phys. **22**, 313 (1968).

23. W.-C. Tan, J.C. Inkson, and G.P. Srivastava, Phys. Rev. B **54**, 14623 (1996).

24. See, e.g., N. Miura, H. Nojiri, and Y. Imanaka, in 22nd Internat. Conf. on the Physics of Semiconductors, edited by D. J. Lockwood (World Scientific, Singapore, 1994), p. 1111.

25. C. Priester, G. Allan, and M. Lannoo, Phys. Rev. B **38**, 9870 (1988).

26. J. E. Sipe and Ed Ghahramani, Phys. Rev. B **48**, 11705 (1993).

27. J.-M. Jancu, R. Scholz, F. Beltram, and F. Bassani, Phys. Rev. B, in press.

28. P. Lautenschlager, M. Garriga, S. Logothetidis, and M. Cardona, Phys. Rev. B **35**, 9174 (1987).

29. A. Di Carlo, S. Pescetelli, M. Pacioti, and P. Lugli, Solid State Commun. **98**, 803 (1996).

30. G. K. Straub and W. A. Harrison, Phys. Rev. B **39**, 10325 (1989).

31. J. A. Majewksi and P. Vogl, in The Structure of Binary Compounds, edited by F.R. de Boer and D. G. Pettifor (Elsevier, New York, 1989), p. 287.

32. H. Bilz and W. Kress, Phonon Dispersion Relation in Insulators (Springer, Berlin, 1979).

33. L. J. Sham, in Dynamical Properties of Solids, edited by G. K. Horton and A. A. Maradudin (Elsevier, New York, 1974), p. 301.

34. P. Vogl, J. Phys. C **11**, 251 (1978).

35. J. Bennetto and D. Vanderbilt, Phys. Rev. B **53**, 15417 (1996).

ORDERING EFFECTS IN DISORDERED METALLIC ALLOYS

A. PASTUREL
Laboratoire de Physique et Modélisation des Milieux Condensés. Maison des Magistères, BP 166 CNRS, 38042 Grenoble-Cedex, FRANCE.

ABSTRACT

This paper reviews recent progress in the theory of the atomic and electronic structure of disordered metallic alloys containing transition metals (TM). Realistic structural models are obtained using molecular dynamics simulations based on pair interatomic forces derived from a tight-binding-bond method. This approach describes quantitatively compositional trends in chemical and topological short-range order in agreement with diffraction experiments.

INTRODUCTION:

It has been known for a long time that order-disorder phenomena affect a majority of the physical properties that most alloys exhibit as a function of temperature, composition and pressure [1]. Depending on the force of the interatomic interactions, the alloys, ordered at low temperature, may or may not disorder before melting. In all cases, if the high temperature phase (solid or liquid) no longer exhibits any long-range order, there still exists a local short-range order that is very important in understanding the properties of these systems. From the point of view of the fundamental theory, there are several aspects of the ordering effect to be studied. The first point is that the interatomic interactions between the atoms depend on the type (A or B) of atom concerned. The problem then arises of knowing how to compute these interactions. In the case of metallic alloys, the characteristics of the bonds involved can be understood by studying their electronic structure. Once the interactions are calculated, there remains the task of determining the thermodynamic functions of alloys. In the past decade, it has been a goal to combine at a high level of accuracy both quantum-mechanical and statistical-thermodynamical contributions to obtain such a theoretical determination of order-disorder phenomena. The main results reported in the literature concern the ordering effects in alloys having a given crystalline structure (see [1] and [2] for a review). In this case, one of the most efficient statistical-mechanics techniques is the cluster-variational method (CVM) which provides a good description of the free energy as a function of the short-range order. Based on a cluster expansion of the configurational free energy, it has been essentially used to deal with solid solutions or ordered phases presenting an extended concentration range. To calculate these cluster interactions from quantum mechanics, two types of approaches have been proposed. The first one is the so-called Connolly-Williams and the closely related renormalized interaction method [3]. These methods use standard ab initio or tight-binding band-structure techniques whose total energies are mapped on to a generalized Ising-like Hamiltonian [2]. In the second approach, the energy of a configurationally disordered state (with no short-range order) as calculated by means of the coherent potential approximation (CPA) is the starting point. The cluster interactions are then calculated using the embedded cluster method [4] or the generalized perturbation method [1]. Either ab initio or tight-binding (TB) Hamiltonians have been used in these approaches [2].

A microscopic theory of the short-range order in topologically disordered alloys like liquid or amorphous alloys seems to be a harder challenge for theoreticians in so far as the

topology of these systems is not known. However it is true that most of the trends in the crystalline structures persist in the liquid state. For instance, in the binary systems, the formation of strictly ordered intermetallic compound is paralleled by similar ordering phenomena in the liquid state which are well-described in terms of the Bhatia-Thornton number-concentration structure factors $S_{NN}(q)$, $S_{Nc}(q)$ and $S_{cc}(q)$ (and their Fourier-transforms, the number-concentration correlation functions $g_{NN}(r)$, $g_{Nc}(r)$ and $g_{cc}(r)$). In an ordered liquid mixture (which is often called ' compound-forming') the concentration-fluctuation structure factor $S_{cc}(q)$ has a peak at small momentum transfers which is very similar to the superlattice peak observed in the diffraction pattern of the corresponding solid alloy [6]. This close analogy suggests that the trends in the liquid structure are driven by electronic effects in a way quite similar to the trends in the crystalline structures.

Thus the study of the short-range order in the liquid or amorphous structures is clearly divided into two distinct parts: (1) the derivation of appropriate expressions for the interatomic forces based on an electronic theory of the chemical bond and (2) the calculation therefrom of the pair correlation functions and of the thermodynamic properties using computer-simulation algorithms like molecular dynamics or more conveniently using some approximate analytical techniques.

This review is organized as to follow this strategy very tightly. First we present a tight-binding bond approach to interatomic forces and describe the calculation of the bond order for the d-d and p(or s)-d interactions. The following section is devoted to the discussion of the short-range order calculated in liquid and amorphous transition metal-based alloys.

TIGHT-BINDING-BOND APPROACH TO INTERATOMIC FORCES:

Bond order:

Here we present only a brief sketch of the basic ingredients for calculating interatomic forces in disordered transition metal (TM) -based alloys. For details we refer to Ref. [7]. Within the tight-binding- bond approach, it has been shown [8] that the variational property of the ground-state energy in the local-density approximation allows a decomposition of the binding energy into a repulsive pair-potential term containing the electrostatic, exchange-correlation, and non-orthogonality contributions to the total energy, and a covalent bond energy measuring the energy gain resulting from the formation of a band with the local density of states (DOS) $n_{i\alpha}(E)$.

$$E_{bond} = \sum_{i\alpha} \int^{E_F} (E - E_{i\alpha}) n_{i\alpha}(E)dE \qquad (1)$$

The bond energy may be broken down in term of contributions from individual pairs of bonds by writing equation (1) as:

$$E_{bond} = \frac{1}{2N} \sum_i \sum_{j(\neq i)} \Phi_{bond}(i,j) \qquad (2)$$

where $\Phi_{bond}(i,j) = \sum_{\alpha,\beta} \Phi^{bond}_{\alpha(i),\beta(j)} = 2 \sum_{\alpha,\beta} H_{i\alpha,j\beta} \Theta_{j\beta,i\alpha}$ $\qquad (3)$

$H_{i\alpha,j\beta}$ is the Slater-Koster bond integral matrix linking the orbitals α, β on site i and j together. Θ is the corresponding bond-order matrix whose elements give the difference between the number of electrons in the bonding $\left|\psi_0^+\right\rangle = \frac{1}{\sqrt{2}}\left|i\alpha + j\beta\right\rangle$ and antibonding states $\left|\psi_0^-\right\rangle = \frac{1}{\sqrt{2}}\left|i\alpha - j\beta\right\rangle$.

254

Bond-order potentials are similar to the embedding potentials in that the bond in a given pair of atoms is considered as embedded in and depending on the local atomic environment. Thus Eq. (3) represents only formally a pair interaction and depends via the bond-order on many-atom effects. This dependence can be explicitly shown by using the recursion method to write the bond-order as an integral over the imaginary part of the difference of two continued fractions [9]:

$$\Theta_{i\alpha,j\beta} = -\frac{1}{\pi} \text{Im} \int^{E_F} \left[G_{00}^+(E) - G_{00}^-(E) \right] dE \qquad (4)$$

with $G^\pm(E) = \left\langle \psi_0^\pm \left| (E-H)^{-1} \right| \psi_0^\pm \right\rangle$

The continued-fraction representation leads to an expansion of the bond-order in terms of many-body cluster interactions. In their study of structural stability in crystalline transition metals, Pettifor and Aoki [10] have shown that the cubic (bcc) versus cubic (fcc) stability depends on the four-body interactions while the cubic (fcc) versus hexagonal (hcp) stability depends on the six-body interactions. For a disordered system however, the detailed structure in the electronic density of states which is the origin of the characteristic variations of the higher moments (and hence of the many-body interactions) is smeared out. The higher moments decay more rapidly and -even more important- we expect the difference in bonding and antibonding moments to be much smaller than in the crystalline states. Hence it should be possible to obtain a much more rapid convergence of the many-body expansion. However, the recursion expansion which is the cornerstone of the bond-order approach to crystalline metals requires the atomic positions to be known whereas in our case the liquid structure is just what we want to derive from the interatomic forces.

Bethe lattice approximation:

To calculate $\Theta_{i\alpha,j\beta}$ and the local DOS $n_{i\alpha}(E)$, we have introduced a Bethe lattice approach which accounts in a realistic way for the shape and variation of the TB-DOS relating to different alloys. In the Bethe lattice reference system, the dependence on the local environment of the bond energy is determined only by the coordination number and the bond length. In this case the bond order can be written as:

$$\Theta_{i\alpha,j\beta} = -\frac{2}{\pi} n_\alpha n_\beta \text{Im} \int^{E_F} \left[T_{\alpha(i),\beta(j)}^2 G_{\alpha(i)} G_{\beta(j)} \right] dE \qquad (5)$$

where n_α, n_β are the degeneracy of subspaces α, β respectively, $T^2_{\alpha(i),\beta(j)}$ is the mean square of the matrix element between a state of subspace α of atom i and a state of subspace β of atom j. $G_{\alpha(i)}$ is the Green's function of atom i in the subspace α as given by the Bethe lattice formalism [11].

$$G_{\alpha(i)}(z) = [z - E_{\alpha(i)} - Z \sum_{\beta(j)} p_{ij} n_\beta T^2_{\alpha(i),\beta(j)} G_{\beta(j)}(z)]^{-1} \qquad (6)$$

Z is the coordination number and p_{ij}, the pair probability. The advantage of Eq.(6) is that it allows us to estimate the different orbital contributions in the bond energy, in particular to treat the effect of hybridization explicitly. Assuming that the bond order is slowly varying function of the bond length [9] we obtain:

$$\Phi^{bond}_{\alpha(i),\beta(j)} = h_{\alpha\beta}(r_{ij}) \Theta_{i\alpha,j\beta} \qquad (7)$$

Alloy effects:

The advantage of the Bethe lattice reference system is that the bond orders in an AB alloy depend very sensitively on the partial DOS's in the alloy. For an alloy with a small difference in the group number of the constituents, the system is close to the common-band limit with only small differences in the partial DOS's. This will lead to bond orders of about equal magnitude, $\Theta_{AA} \sim \Theta_{AB} \sim \Theta_{BB}$, comparable strength of the interatomic interactions and finally a random distribution of the two atomic species in the melt. For an alloy with a large difference in the group numbers the electronic DOS is close to the split-band limit, with the two partial DOS's concentrated in the lower and upper parts of the band. This results in strong attractive interactions in AB forces and in the liquid and glassy structure this is reflected in a strong chemical and topological short-range order.

In principle the partial DOS's and the bond order depend on the degree of chemical order assumed to exist on a Bethe lattice. This influence can be seen from Eq. (6). Indeed the pair probabilities p_{KL} are related to the Cowley short-range-order parameter σ through (x_A and x_B are the concentrations of A and B atoms):

$$p_{AA} = x_A + x_B \sigma \qquad p_{AB} = x_B(1-\sigma)$$
$$p_{BA} = x_A(1-\sigma) \qquad p_{BB} = x_B + x_A \sigma \qquad (8)$$

$\sigma > 0$ corresponds to clustering of unlike atoms, $\sigma < 0$ to a tendency to heterocoordination.
An ordering bond energy is then defined as the energy relative to the random configuration:

$$\Delta E_{bond}^{ord}(\sigma) = E_{bond}(\sigma) - E_{bond}(0) \qquad (9)$$

On the other hand, in terms of pair interactions, $\Delta E_{ord}^{bond}(\sigma)$ is given by using the pair probabilities $P_{KL} = x_K p_{KL}$:

$$E_{bond}(\sigma) = \frac{1}{2N} \sum_{K,L} \sum_i \sum_{j(j \neq i)} P_{KL}(r_{ij}) \Phi_{bond,KL}(r_{ij}) \qquad (10)$$

$$= E_{bond}(0) + \frac{1}{4}\sigma \sum_i \sum_{j(j \neq i)} \left\{ x_A x_B \left[\Phi_{bond,AA}(r_{ij}) + \Phi_{bond,BB}(r_{ij}) - 2\Phi_{bond,AB}(r_{ij}) \right] \right\} \qquad (11)$$

We find that in a nearest-neighbor approximation, the bond contribution to the ordering energy is given, to first order in σ, by the ordering potential:

$$\Phi_{cc}(r) = x_A x_B [\Phi_{bond,AA}(r) + \Phi_{bond,BB}(r) - 2\Phi_{bond,AB}(r)] \qquad (12)$$

evaluated in the random reference configuration.
A first requirement is that the pair potentials derived from the random configuration yield a reasonably accurate description of ordering phenomena, that means the ordering energies calculated via (9) and (11) display similar values. This condition gives the capacity of the bond-order potentials to describe the short-range order in a disordered structure if they are coupled with molecular dynamics simulations or with thermodynamic variational techniques.

SHORT-RANGE ORDER IN TM-BASED ALLOYS:

Tight-binding Hamiltonian:

The calculation of the bond energies of the TM-based alloys include the five d-like and one s-like orbitals for TM and the three p-like and one s-like orbitals for p-metals (like Al or B). The hopping energies $h_{i\alpha,j\beta}$ between like species were evaluated from Harrison's solid-state table [12]. The distance dependence of the hopping integrals is evaluated according to Harrison's power law ($h_{\alpha\beta}(r_{ij}) = \dfrac{h_{\alpha\beta}(1)}{r_{ij}^n}$), n=2 for s and p orbitals, n=5 for d orbitals. For the on-site energies, we use the atomic on-site energies given by Herman and Skillman [13]; for transition metals, the values were calculated for a transition atom with the $d^{n+1}s$ configuration.

The repulsive part of the binding energy is assumed to be given by a sum over pair potentials:

$$E_{rep} = \frac{1}{2N} \sum_i \sum_{j(j \neq i)} \Phi^{rep}(r_{ij}) \tag{13}$$

with $\Phi^{rep}(r_{ij}) = \sum_\alpha \Phi^{rep}_{\alpha,\alpha}(r_{ij})$

For s-s and p-p interactions, the repulsive pair interactions may be modeled as $\Phi^{rep}_{\alpha,\alpha}(r_{ij}) = \dfrac{C_{\alpha,\alpha}}{r_{ij}^4}$ while for d-d interactions, a stronger power-law dependence $\dfrac{C_{\alpha,\alpha}}{r_{ij}^{10}}$ is chosen . $C_{\alpha,\alpha}$ (α=s, p, or d orbitals) are the only parameters of the model; they are determined from the knowledge of the experimental atomic value and bulk modulus of the pure metals. This usual dependence for both hopping integrals and repulsive terms has been modified using the rescaling method proposed by Goodwin et al. [14]. This method is known for generating improved TB parameters which are both transferable and suitable for extensive molecular dynamics simulations.

To treat the alloying effect, i.e. the AB interactions, we use geometrical average for hopping integrals as well as for the parameters entering the repulsive part. Consequently, no parameters are introduced to describe the alloy properties. The shift between the on-site energies of the two metals is determined self-consistently by imposing local charge neutrality, a reasonable approximation for TM based alloys, where charge transfers are known to be small.

Table I shows all the parameters entering the interatomic forces for the alloys which are discussed in the present paper.

A possible question concerns the validity of the Bethe lattice approach to generate a correct short-range order. Therefore we have calculated effective pair interactions between first nearest neighbors, i.e. V_1, for the fcc $Ni_{50}Ti_{50}$ alloy and we have compared this result with the ones obtained from the generalized perturbation method (GPM) using the same tight binding hamiltonian or with first-principles based results. The Bethe lattice gives a value of 3.7 mRy/atom very close to the one obtained from GPM, i.e. 4.0 mRy/atom; it is also in fair agreement with the one derived using Connolly-Williams inversion scheme with FPLMTO calculations [15], 4.4 mRy/atom or with the one obtained from KKR-GPM-CPA approach [16], 1.86 mRy/atom. Another instructive comparison can be done for the bcc $Al_{50}Ni_{50}$ alloy. The Bethe lattice value is 5.1 mRy/atom compared to 7.6 mRy/atom extracted from FPLMTO calculations or to 6.2 mRy/atom obtained from KKR-GPM-CPA approach.

Table I: Tight-binding and repulsive parameters for Al and TM (in eV)

	ssσ	ddσ	ddπ	ddδ	sdσ	ppσ	ppπ	spσ	C_{ss}	C_{dd}	C_{pp}
Al	-0.52					1.20	-0.31	0.71	0.30		0.30
Cu	-1.32	-0.42	0.21	0	-0.92				0.08	0.23	
Ni	-1.42	-0.56	0.28	0	-1.11				0.53	0.27	
Co	-1.36	-0.60	0.30	0	-1.15				0.78	0.30	
Fe	-1.30	-0.66	0.33	0	-1.15				1.18	0.32	
Mn	-1.33	-0.65	0.32	0	-1.22				1.60	0.14	
Ti	-0.99	-0.84	0.42	0	-1.13				0.33	0.32	
Y	-0.67	-0.96	0.48	0	-0.99				1.15	0.22	

Thermodynamic variational methods versus Molecular Dynamics simulations:

For many liquid alloys, a reasonably accurate description of the structures and the thermodynamic properties may be derived from the thermodynamic variational approach based on the Gibbs-Bogoljubov inequality, provided that a sufficiently realistic system is available , as the Hard-Sphere Yukawa (HSY) reference system [17]. Within this model, the structural manifestations of ordering are modeled by a mixture of hard spheres all having the same diameter but opposite charge (while respecting the overall charge neutrality condition) which interact by a screened Coulomb (Yukawa) potential. In this system, the description of the atomic configuration requires three parameters which are the diameter of the hard sphere d, the strength of the ordering potential at hard contact ε, and a screening constant κ. The reference system provides expressions for the excess entropy and the partial pair correlation functions. The three parameters are then obtained by minimizing the free energy given by:

$$F(d,\varepsilon,\kappa) = \frac{3}{2}k_B T + E_{bind}(d,\varepsilon,\kappa) - T[S_{HS}(d) + S_{ord}(d,\varepsilon,\kappa)] \qquad (14)$$

E_{bind} (d,ε,κ) means that the pair probabilities entering the expression of the binding energies (as in Eq. (10)) are replaced by the partial pair correlation functions as given by the HSY reference system.
The weakness of this approach is the assumption of equal diameter for all hard spheres which means that this reference system is not able to treat the influence of size effects on short-range order. To check the results of this variational procedure it is then important to perform molecular dynamics simulations. We have performed microcanonical molecular dynamics (MD) simulations based on integrating the equations of motion in a velocity form of the Verlet algorithm with a time increment of 10^{-15} s. A 1332-particle cluster with periodic boundary conditions and the lattice parameter of the MD cell is expanded to the required liquid density. The liquid is 'heated' by raising the initial temperature and subsequently scaling it down to the required value. Typical simulations run up to 3-4 10^4 steps. Pair correlation functions are based on averages over 40 independent configurations taken at intervals of 100 time steps. For a production of the glassy phase, the liquid alloy was compressed isothermally to the density of the glass and reequilibrated. Subsequently, the liquid is quenched to room temperature (quenching rate 10^{14} Ks^{-1}) at constant volume. After quenching, the system is equilibrated for 2000 time steps, and finally 4000 time steps and 40 independent configurations are used for calculating pair correlation functions.

Short-range order in TM-TM alloys:

TM-TM glasses can be made over a wide range of compositions and for systems with a strong (ex Ni-Y) to weak (Fe-Y) chemical short-range order. Diffraction studies indicate a strongly varying degree of chemical as well as topological order [18]. Ni-based metallic glasses are particularly interesting because only for the Ni-alloys accurate partial correlation functions and structure factors are available from isotope-substitution experiments.

We first present some results for amorphous $Ni_{40}Ti_{60}$ alloy. This system is characterized by large differences in the number of d-electrons and in the atomic d-electron eigenvalues. The electronic density of states is close to the split-band limit as seen in Fig. 1a and strongly non additive pair potentials (Fig. 1b) are expected.

Figure 1: (a) DOS for the $Ni_{40}Ti_{60}$ alloy: (-----) total; (.......) Ni-DOS; (------) Ti-DOS. (b) Tight-binding-bond potentials (eV): (-----) Ni-Ni interaction; (-----) Ni-Ti interaction; (......) Ti-Ti interaction.
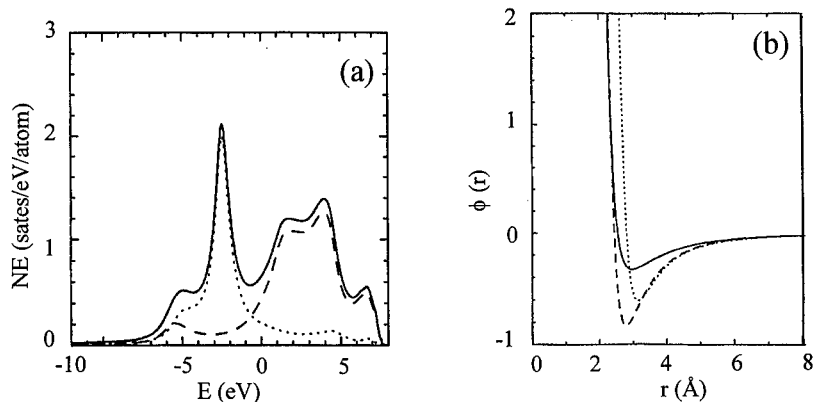
Fig. 2 displays the partial concentration structure factor $S_{cc}(q)$. $Ni_{40}Ti_{60}$ alloy shows a strong CSRO as indicated by the significant oscillations of $S_{cc}(q)$ up to 10 A^{-1}. Due to the weak size effects, the thermodynamic variational method and the molecular dynamics simulations give similar results.

In Fig. 3 we present the calculated ordering energies and ordering entropies for the Ni-Ti alloys as a function of the composition. We can see that these evolutions are very peculiar, the two quantities display a minimum around $x_{Ni}=0.8$.

Figure 2 : Comparison between the $S_{cc}(q)$ curves calculated by :(a) molecular dynamics
simulation; (b) variational procedure.
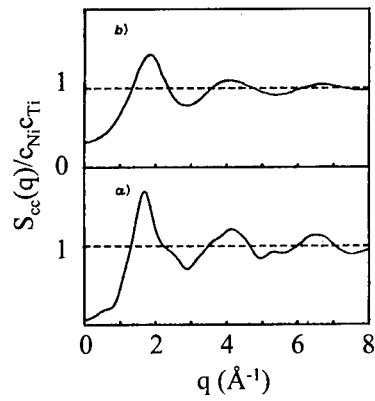
Figure 3: Ordering energies and ordering entropies of amorphous
Ni_xTi_{1-x} alloys.
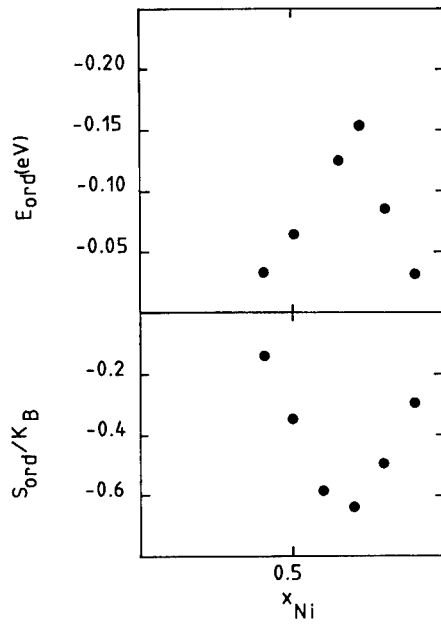
As already mentioned, the calculated DOS of the NiTi alloys is characterized by a two-peaked
structure and in each peak the DOS is largely dominated by the contribution of one species
since the mixing is weak between the two bands. On the contrary, between the two peaks, the

since the mixing is weak between the two bands. On the contrary, between the two peaks, the heteroatomic coupling is important. The effect of CSRO is to diminish the number of homoatomic pairs and to increase the number of heteroatomic pairs. The consequences on the shape of the alloy DOS is that the CSRO diminishes the bandwidths of the two subbands but increases their coupling, which is important between the two peaks. Thus, if the band filling is such that the Fermi level is in one of the two peaks of the DOS, the ordered state is energetically unfavorable since the bandwidths are reduced. In contrast, if the Fermi level resides between the two peaks, the bond energy is sensitive to the increase in the coupling between the two subbands and CSRO is energetically more favorable. This simple argument allows us to understand why E_{ord} presents a minimum at the Ni-rich end, the electronic configuration being the most favorable one for the occurrence of CSRO.

In a second step we present results for a series of $Y_{35}TM_{65}$ alloys with TM=Mn, Fe, Co, Ni, Cu. With a decreasing difference in the group number, the electronic DOS of the alloys gradually moves from a split-band regime in YNi alloy to a nearly common-band DOS in YMn alloy.

Figure 4: Tight-binding-bond potentials (eV); (------) TM-TM interaction; (------)Y-TM interaction; (.........) Y-Y interaction (a) $Y_{35}Ni_{65}$; (b) $Y_{35}Fe_{65}$.
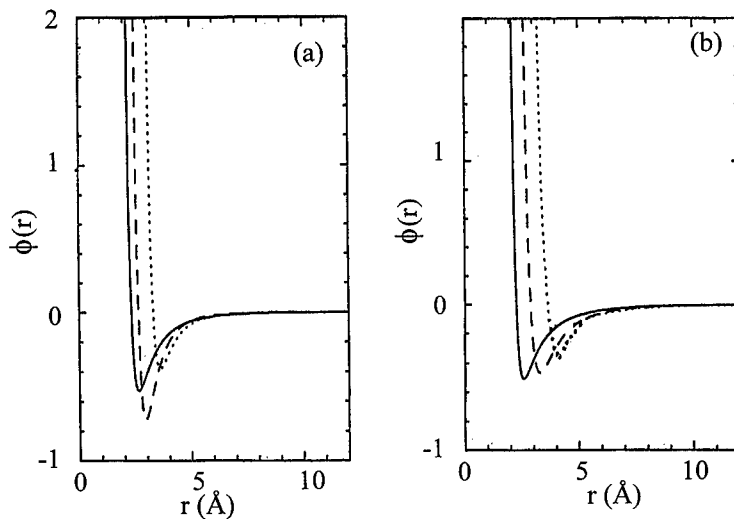
As for the NiTi alloy, the interaction for YNi pairs is larger than the interactions for NiNi and YY pairs while for the YMn alloy, all three pair interactions are of comparable of magnitude (see Fig. 4) .The consequences on the CSRO can be seen from the calculated Cowley short-range-order parameter: its value is continuously decreasing from 0.005 to -0.14 when the number of d electrons of TM increases.

The analysis of these results demonstrates a clear trend from strong to moderate chemical order with decreasing difference in the group number and with decreasing Ni-content within a given system. Using molecular dynamics simulations allows also to analyze the topological short-range order in liquid or glassy TM-TM alloys. Even if this discussion is beyond the scope of this paper it is important to emphasize that the bond-order approach

predicts a gradual change from a trigonal-prismatic to a polyhedral topological short-range order within the series Y-Cu, Y-Ni, Y-Co, Y-Fe, Y-Mn as investigated from the total and partial bond-angle distribution functions [19]. This change may be traced back to the variation of the interatomic forces and of the electronic structure. A pronounced non-additivity of the pair interactions and a strong short-range interaction between unlike atoms lead to the formation of trigonal-prismatic clusters. As the non-additivity is reduced, icosahedral clusters are preferred.

Short range order in TM-Al alloys:

Alloys formed between transition metals and elements with only s and p electrons in their valence shells display high negative formation energies. Such a behavior can be understood from a strong mixing between the d states of the transition metal and the sp states of the polyvalent metal: the resulting formation of bonding and antibonding hybridization states is the dominant factor with the occurrence of a pseudogap at the top of the d band [11]. The occurrence of CSRO is then related to the occurrence of this pseudogap and the location of the Fermi level in it. It is then interesting to know if such a behavior persists in the liquid TM-Al alloys. During the last few years, short-range order in TM-Al liquid alloys has been extensively studied through an accurate determination of the partial pair correlation functions by neutron diffraction [20, 21]. In this contribution, I have focused attention on the study of topological and chemical SRO in $Al_{80}Mn_{20}$ and $Al_{80}Ni_{20}$. The first distances in both Al-Al and TM-TM distributions differ significantly in the two alloys, pointing out distinct topological ordering. Then it is important to know if the bond-order potentials are able to reproduce such differences.

Figure 5: Tight-binding-bond potentials (eV): (———)TM-TM interaction; (------)TM-Al interaction; (.........)Al-Al interaction. (a) $Al_{80}Ni_{20}$; (b) $Al_{80}Mn_{20}$.
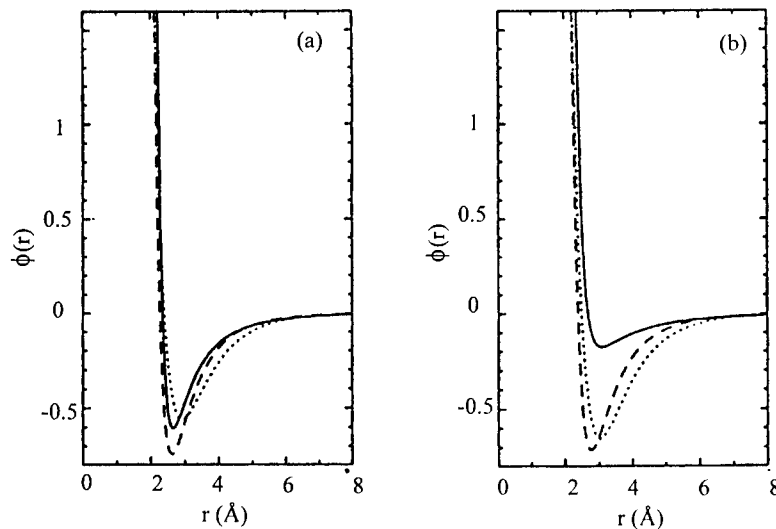
Fig. 5 shows a strong interaction between pairs of unlike atoms at short-bond distances. This is the consequence of a strong hybrization between p states of Al and d states of TM. The Mn-Mn interaction is found to be very small in comparison with Ni-Ni interaction and explains the two different topologies obtained for both alloys. Indeed we find that TM-TM distance is equal to 2.55 A in $Al_{80}Ni_{20}$ while it is equal to 2.86 A in $Al_{80}Mn_{20}$. This difference cannot be attributed to the very small atomic size difference between Ni and Mn atoms and suggests different local arrangements. This is also supported by the first Al-Al distances significantly shorter than Mn-Mn contacts in $Al_{80}Mn_{20}$ and greater than Ni-Ni contacts in $Al_{80}Ni_{20}$. Experimental data give 2.82 and 2.75 A for $Al_{80}Ni_{20}$ and $Al_{80}Mn_{20}$ respectively, while theoretical analysis gives 2.80 and 2.76 A. The distributions of the first heteroatomic pairs are centered at the same position (2.54 A for experiments and 2.55 A for calculations), which corresponds to short distances in comparison with TM-TM and Al-Al contacts. The fact that these interatomic forces are able to reproduce so different topological short-range orders in both alloys lead us to believe that the dependence of the pair interactions on the bond order determined by the strenght of the pd hybridization is a major improvement to understand the structure of the liquid transition metal-aluminum alloys.

CONCLUSION:

We have shown that it is now possible to present a coherent microscopic theory of the structural and electronic properties of disordered transition metal alloys based on quantum-mechanically derived interatomic forces. The pronounced non additivity of the pair interactions and a strong short-range interaction between unlike atoms is due to the strong covalent forces characteristic for alloys with strong d-d or p-d hybridization. This strong hybridization can be related to the occurrence of a pseudogap in the DOS alloy and the location of the Fermi level in it. There are certain indications in the measured correlation functions that the local topology is somewhat more sharply defined in the real material than in our models. To eliminate these remaining small differences it will be necessary to go beyond the pair-potential approximation. However, both the quantum-mechanical calculation of angular-dependent forces by a moment expansion of the bond order and their use in MD simulations constitute a considerable complication.

REFERENCES:

1. F. Ducastelle, in Order and Phase Stability in Alloys (Cohesion and Structures Series, Vol.3) eds F.R. de Boer and D.G. Pettifor, North-Holland, Amsterdam, 1991.

2. P.E.A. Turchi, in Electronic Theories of Alloy Phase Stability. Intermetallic Compounds Vol.1, Principles, eds. J.H. Westbrook and P.L. Fleischer, John Wiley and Sons Ltd, 1994.

3. A. Zunger, in Statics and Dynamics of Alloy Phase transformations, eds. P.E.A Turchi, A. Gonis, NATO-ASI Series B, 319, Plenum press New York 1994, p. 361.

4. A. Gonis, X.G. Zhang, A.J. Freeman, P.E.A Turchi, G.P. Stocks, and D.M. Nicholson, Phys. Rev. B 36, 4630 (1987).

5. F. Ducastelle and F. Gautier, J. Phys. F **6**, 2039 (1976).

6. J. Hafner, G. Kahl and A. Pasturel, in Amorphous and Liquid Materials, eds. E. Lüscher and J. Jacucci, NATO-ASI Series E, **118**, Plenum Press New York 1987, p. 164.

7. L. Do Phuong, D. Nguyen Manh and A. Pasturel, Phys. Rev. Lett. **71**, 372 (1993).

8. A.P. Sutton, M.W. Finnis, D.G. Pettifor and Y.Ohta, J. Phys. C **21**, 35 (1988).

9. D.G. Pettifor, Phys.Rev. Lett. **63**, 2480 (1989).

10. D.G. Pettifor and M. Aoki, Philos. Trans. R. Soc. London, Ser.A **334**, 439 (1991).

11. D. Mayou, D. Nguyen manh, A. Pasturel, and F. Cyrot-Lackmann, Phys. Rev. B **33**, 3384 (1986).

12. W.A. Harrison, in Electronic Structure and the Properties of Solids, (Freeman, San Francisco, 1980).

13. F. Herman and S. Skillman, in Atomic Structure Calculations (Prentice-Hall, Englewwod Cliffs, NJ, 1965).

14. L. Goodwin, A.J. Skinner and D.G. Pettifor, Europhys. Lett. **9**, 701 (1989).

15. A. Pasturel, C. Colinet, D. Nguyen Manh, A.T. Paxton and M. van Schilfgaarde, Phys. Rev.B **52**, 15176 (1995).

16. M. Sluiter, P.E.A. Turchi, F.J. Pinski and G.M. Stocks, J. of Phase Equilibria **13**, 605 (1992).

17. A. Pasturel and J. Hafner, Phys. Rev. B **34**, 8357 (1986).

18. M. Maret and A. Pasturel, J. de Physique **48**, 1541 (1987).

19. Ch. Hausleitner and J. Hafner, Phys. Rev.B **45**, 115 (1992).

20. M. Maret, A. Pasturel, J.M. Dubois and P. Chieux, J. Phys. (Paris) **50**, 295 (1989).

21. M. Maret, T. Pomme, A. Pasturel and P. Chieux, Phys. Rev. B **42**, 1598 (1990).

SELF-CONSISTENT TIGHT-BINDING APPROXIMATION INCLUDING POLARISABLE IONS

M.W. FINNIS, A.T. PAXTON, M. METHFESSEL[†] and M van SCHILFGAARDE[‡]
Department of Pure and Applied Physics, Queen's University, Belfast, BT7 1NN, UK
[†]Institute for Semiconductor Physics, Walter-Korsing-Str. 2
    D–15230, Frankfurt (Oder), Germany
[‡]SRI International, Menlo Park, California 94025, USA

ABSTRACT

Until recently, tight-binding has been applied to either covalent or metallic solid state systems, or charge transfer treated in a simple point charge framework. We present a self-consistent tight-binding model which, for the first time, includes electrostatic ion polarisability and crystal field splitting. The tight-binding eigenvectors are used to construct multipole moments of the ionic charges which are used to obtain angular momentum components of the electrostatic potential in structure constant expansions. Our first test of the model is to study the phase stability in zirconia; in particular the instability of the fluorite phase to a spontaneous symmetry breaking, and its interpretation in terms of band effects and ion polarisability. This new formalism opens up the way to apply the tight-binding approximation to problems in which *polarisation* of atomic charges is important, for example oxides and other ceramic materials and surfaces of metals.

INTRODUCTION

Although the tight-binding approximation was originally conceived to deal with homopolar covalently bonded solids, an extension to include charge transfer in alloys and heteropolar semiconductors was proposed by Falicov and by Harrison some 15–20 years ago [1,2]. This self-consistent solution was extended and used very effectively by Majewski and Vogl to describe the energetics of ionically *sp*-bonded compounds [3]. A self-consistent scheme for metals has also been used [4], which in its simplest form assumes local charge neutrality [5] and leads to the tight-binding bond model [6]. Recently some questions have arisen that have motivated us to extend the self-consistent tight-binding approximation with charge transfer to include the effects of polarisation of the atomic charges. Such observations include,

1. The electronic structure of a low coverage of Nb on the (0001) α-Al_2O_3 surface has been observed to have the character of a single non degenerate d-orbital. We found it impossible to describe this with conventional orthogonal or non orthogonal tight-binding, in which the on site hamiltonian matrix elements are constrained to be the same for each angular momentum (*ie*, crystal field splitting is ignored).
2. There is a question whether surface dipole barriers and hence work functions could be calculated in the tight-binding approximation.
3. It is thought that the spontaneous distortion of cubic zirconia into the tetragonal structure is driven by the quadrupole polarisability of the oxygen atoms [7].
4. One would like to include crystal field terms as a natural consequence of the non spherical Madelung potential, rather than by direct parameterisation as has been done previously [8].

THEORY

Traditional self-consistent tight-binding

In order to make our new approach clear we begin with a brief reminder of Harrison's scheme which will also serve to establish our notation. In non self-consistent tight-binding, one supposes the existence of a hamiltonian $\mathcal{H}_0$ whose matrix elements between localised (not necessarily orthogonal) basis functions are $\langle \mathbf{R}'L' | \mathcal{H}_0 | \mathbf{R}L \rangle$. Although the basis functions are never specified, without loss of generality one may write them as a radial part times an angular part:

$$|\mathbf{R}L\rangle = f_{\mathbf{R}\ell}(|\mathbf{r} - \mathbf{R}|) Y_L(\mathbf{r} - \mathbf{R}). \tag{1}$$

Here we use $\mathbf{R}$, $\mathbf{R}'$ to denote the position of atomic sites with respect to some origin. To simplify the equations, we will frequently suppress the site index, in which case one can take it that we are referring to an atom at the origin and $\mathbf{r}$ is a small vector in its neighbourhood. L is a composite angular momentum index $L = \{\ell m\}$ and Y_L is a real spherical harmonic [9].

In the traditional self-consistent approach, having solved Schrödinger's equation, the eigenvectors are used to determine to what extent charge has been redistributed among the atoms. There will be a charge

$$q = Q_0 Y_0 = e \sum_\ell N_\ell$$

on each site, where Q_0 is the monopole moment of the charge, N_ℓ is the Mulliken population of angular momentum ℓ (s, p and d) and e is the charge on the electron. The hamiltonian is adjusted self-consistently to become $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'$ where $\mathcal{H}'$ *is diagonal* in $\mathbf{R}L$:

$$\mathcal{H}'_{\mathbf{R}L\mathbf{R}L} = eV_{\mathbf{R}0} + U_{\mathbf{R}\ell} N_{\mathbf{R}\ell}. \tag{2}$$

Here, $V_{\mathbf{R}0}$ is the $\ell = 0$ component of the electrostatic potential at site $\mathbf{R}$ due to the charges at all other sites, and $U_{\mathbf{R}\ell}$, sometimes called the "Hubbard U" is a positive energy which resists the accumulation of charge. The physical meaning of U in this context is discussed in detail by Harrison [2]. The benefit of this model which goes beyond non self-consistent tight-binding is that one can treat mixed ionic–covalent solids with real success [3].

Self-consistent tight-binding with multipoles

What is missing, and what is added in our new formulation here, is any adjustment of the *off-diagonal* on-site matrix elements of $\mathcal{H}'$. These matrix elements couple orbitals of different angular momenta, ℓ and m, on the same site. One effect this can have is to break the degeneracy of the s, p and d orbital energies, and thereby induce effects such as crystal field splitting or phenomena such as that mentioned as item 1 in the introduction. Furthermore, these new matrix elements lead to *polarisability* of the ionic valence electrons. One could then study, for example, polarisable anion theories of oxide structure and lattice dynamics, or phenomena such as lattice instability and ferroelectricity. What couple these orbitals on the same site are the higher angular momentum components of the electrostatic potential. In fact we may regard V_0 in (2) as just the coefficient of the first term in an expansion in spherical waves of the electrostatic potential at a point $\mathbf{r}$ near an atom at the origin:

$$V(\mathbf{r}) = \sum_L V_L r^\ell Y_L(\mathbf{r}). \tag{3}$$

For $\ell = 1$ and $\ell = 2$ respectively, V_L are proportional to the electric fields and field gradients at the nucleus [9]. It is the $\ell = 4$ components of V_L that are responsible for the splitting of the d levels by cubic, octahedral or tetrahedral crystal fields in transition metal oxides [10].

In a similar vein we would expect the charge on each atom to have not only $l = 0$ components as in (2), but higher multipole moments, which would be

$$Q_L = e \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})r^\ell Y_L(\mathbf{r}) + q^c Y_0 \delta_{L0} \tag{4}$$
$$= Q_L^e + Q^c \delta_{L0}$$

if we knew the electronic charge density $e\rho(\mathbf{r})$ associated with each atom. We have included here $Q^c = q^c Y_0$ which is the monopole moment of the core charge q^c, and in the second line have separated the total charge into a part from the valence electrons and a part from the ion cores. If we assume the charges are sufficiently localised about their atomic sites, we can find a relation for the components of the electrostatic potential $V_{\mathbf{R}L}$ on one site, due the multipole moments of the charges on all other sites. Using Poisson's equation and an expansion in the structure constants of LMTO-ASA theory, we find

$$V_{\mathbf{R}L} = 4\pi \sum_{\mathbf{R}' \neq \mathbf{R}} \sum_{L'} \tilde{B}_{LL'}(\mathbf{R}' - \mathbf{R}) Q_{\mathbf{R}'L'}.$$

The structure constants are [11,12]

$$\tilde{B}_{LL'}(\mathbf{R}) = 4\pi \sum_{L''}{}' \frac{(2\ell''-1)!!}{(2\ell+1)!!(2\ell'+1)!!} \left(\frac{1}{R^{\ell''+1}} \right) (-1)^{\ell'} Y_{L''}(\mathbf{R}) C_{L''L'L}$$

(with the sum over L'' restricted to values for which $\ell'' = \ell + \ell'$) and

$$C_{L''L'L} = \int \mathrm{d}\Omega \, Y_{L''} Y_{L'} Y_L \tag{5}$$

are Gaunt coefficients for real spherical harmonics. The sum over $\mathbf{R}'$ is done by the Ewald method.

<u>From eigenvectors to total energy, and some new parameters</u>

Although the charge density is not calculated in an empirical tight-binding model, we can nevertheless obtain the multipole moments (4) in the following way. Equation (4) demonstrates that Q_L^e is the expectation value of a certain function of the position operator $\hat{\mathbf{r}}$, namely [9]

$$\hat{Q}_L^e = e \, \hat{r}^\ell \, Y_L(\hat{\mathbf{r}}). \tag{6}$$

Now, a tight binding calculation normally provides eigenvectors for each band index n and wavevector $\mathbf{k}$, $c_{\mathbf{R}L}^{n\mathbf{k}}$, from which one obtains the charges on each site. We can use these to obtain also the multipole moments as the expectation value of the operator $\hat{Q}_L^e$:

$$Q_L^e = \sum_{L'L''} \sum_{\substack{\mathrm{occ.}\\ n\mathbf{k}}} \bar{c}_{L'}^{n\mathbf{k}} c_{L''}^{n\mathbf{k}} \langle L' | \hat{Q}_L^e | L'' \rangle. \tag{7a}$$

We can now use (1), (5) and (6) to express the matrix elements of $\hat{Q}_L^e$ as

$$\langle L' | \hat{Q}_L^e | L'' \rangle = e \, \Delta_{\ell'\ell''\ell} \, C_{L'L''L}. \tag{7b}$$

267

Included in (7) are important quantities $\Delta_{\ell'\ell''\ell}$ which will be introduced as new parameters of the empirical tight binding model:

$$\Delta_{\ell'\ell''\ell} = \int r^2 dr f_{\ell'}(r) f_{\ell''}(r) r^{\ell}.$$

In the case that $\ell' = \ell''$ these are parameters $\langle r^{\ell} \rangle$ already familiar from crystal field theory [10].

There is only a limited number of new "Δ-parameters" due to selection rules and symmetries of the $C_{L'L''L}$ and the $\Delta_{\ell'\ell''\ell}$. Also $\Delta_{\ell\ell'0} = \Delta_{\ell\ell0}\delta_{\ell\ell'}$ and is determined by the normalisation of (1). The new parameters that will have to be determined (for a tight binding basis up to $\ell = 2$) are

$$\Delta_{011} = \Delta_{101} \equiv \Delta_{spp}$$
$$\Delta_{112} \equiv \Delta_{ppd}$$
$$\Delta_{022} = \Delta_{202} \equiv \Delta_{sdd}$$
$$\Delta_{121} = \Delta_{211} \equiv \Delta_{pdp}$$
$$\Delta_{222} \equiv \Delta_{ddd}$$
$$\Delta_{123} = \Delta_{213} \equiv \Delta_{pdf}$$
$$\Delta_{224} \equiv \Delta_{ddg}.$$

The physical meaning of these parameters in the context of equation (7) is as follows. Occupied orbitals of angular momentum (s, p or d-type) ℓ' and ℓ'' combine to produce a multipole moment on the same site of angular momentum ℓ. The strength of the multipole is proportional to the size of $\Delta_{\ell'\ell''\ell}$ and the Gaunt coefficients dictate the selection rules.

The next step is to construct matrix elements of $\mathcal{H}'$. This is done in much the same way as in making matrix elements of $\hat{Q}_L^c$ since $V(\mathbf{r})$ is also expanded in r^{ℓ}. (Compare equations (3) and (6).) We therefore obtain

$$\mathcal{H}'_{\mathbf{R}L'\mathbf{R}L''} = e \sum_L V_{\mathbf{R}L} \Delta_{\ell'\ell''\ell} C_{L'L''L} + U_{\mathbf{R}\ell'} N_{\mathbf{R}\ell'} \delta_{L'L''} \tag{8}$$

in analogy with equation (2), in which, perhaps not surprisingly, the $\Delta_{\ell'\ell''\ell}$ appear for a second time. In this context their physical interpretation is that multipoles on all other sites have induced a potential on a particular site. The ℓ-component of that potential then causes a coupling in $\mathcal{H}'$ between ℓ' and ℓ'' orbitals on that site. The quantity $\Delta_{\ell'\ell''\ell}$ describes the strength of that coupling. This is essentially the same explanation as in the previous paragraph but viewed from different standpoint. Drawing again on the example of crystal field splitting in transition metal oxides, the $\ell = 4$ component of the potential on the transition metal nucleus, due to the surrounding oxygen neighbours, causes a coupling between the d-orbitals which breaks their degeneracy. The strength of this coupling (called the cubic field splitting, Δ [10]) is proportional to the integral $\Delta_{ddg} = \langle r^4 \rangle$.

We can complete our model, finally, by working out the total energy. In conventional non self-consistent tight-binding, the total energy, E_T, contains two terms: the sum of one electron eigenvalues, E_{band} and a repulsive pair-wise potential energy E_{pair} [13]. In our case since we are including the total (core plus valence) charge, and since the electron–electron interaction
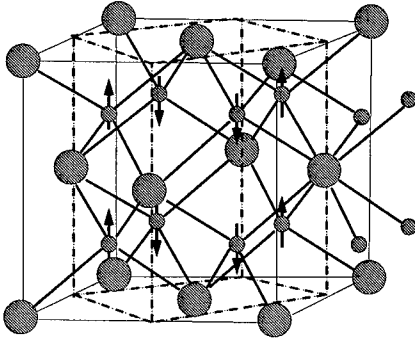
FIG 1: Atomic positions in the fluorite and tetragonal modifications of ZrO_2. Large circles are Zr atoms and small circles oxygen. Arrows represent the displacements of the columns of oxygen atoms that cause the spontaneous transition from cubic to tetragonal ZrO_2. A broken line outlines a unit cell of tetragonal ZrO_2. In the present model calculations we retain the cubic lattice constant so that the axial ratio of the tetragonal phase is constrained at $\sqrt{2}$.
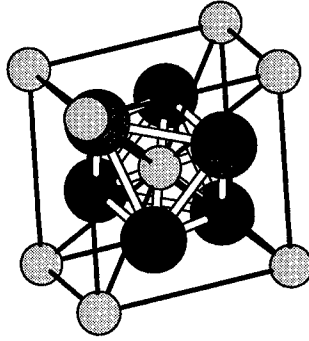
FIG 2: Atomic structure of the rutile phase. Here the small atoms are Zr. Note how whereas in fluorite (as seen on the right of fig 1) the transition metal atom is in the centre of a cube of anions, in rutile the transition metal atom is in an octahedral crystal field, which reverses the order of the t_2 and e sub band energy levels (see fig 4).

will be double counted in E_{band}, we will need to add two further terms, the ion-ion repulsive electrostatic energy, E_{ii}, and the double counting correction $-E_{\text{DC}}$. We find

$$
\begin{aligned}
E_{\text{T}} &= E_{\text{band}} + E_{\text{pair}} + E_{\text{ii}} - E_{\text{DC}} \\
&= \langle \mathcal{H} \rangle + E_{\text{pair}} + \frac{1}{2} e \sum_{\mathbf{R}} Q_{\mathbf{R}}^{\text{c}} V_{\mathbf{R}0} - \frac{1}{2} e \sum_{\mathbf{R}L} Q_{\mathbf{R}L}^{\text{e}} V_{\mathbf{R}L} - \frac{1}{2} \sum_{\mathbf{R}\ell} U_{\mathbf{R}\ell} N_{\mathbf{R}\ell}^{2}.
\end{aligned}
$$

Using $\langle \cdots \rangle$ we have written the usual band energy as the expectation value with respect to the self consistent one electron eigenvectors of the hamiltonian $\mathcal{H}$. We may obtain an exact derivative of E_{T} in order to calculate the interatomic forces. This is a generalisation of the expression due to Sutton *et al.* [6] since we include a finite U. Omitting small dipole–dipole and dipole–higher-multipole terms (which can be written down explicitly) we find, for the z-component of the force on an atom at the origin,

$$
F_z = -2 \left\langle \frac{\partial \mathcal{H}_0}{\partial z} \right\rangle - \frac{\partial E_{\text{pair}}}{\partial z} + \sqrt{\frac{3}{4\pi}} \, Q_0 V_{10}.
$$

Note the first two terms are exactly the force used in non self consistent tight-binding [6,13], except the eigenvectors are self consistent. The final term accounts for the double counting (the "U" terms cancel) and V_{10} as in (3) is the $\ell = 1$ component of the potential ($m = 0$ for the z-component of the force) and Q_0 is the total charge as in (4).

APPLICATION TO PHASE STABILITY IN ZIRCONIA

In order to demonstrate the model in an example, we address the question mentioned in item 3 in the introduction, namely the phase stability in zirconia, ZrO_2. The zero temperature, ground state structure is a complex monoclinic phase. Below the melting point ZrO_2 has the fluorite structure, illustrated in fig 1; but at intermediate temperatures this structure transforms to a tetragonal phase in a spontaneous symmetry breaking also illustrated in fig 1. We will concern ourselves here with model calculations aimed at predicting this symmetry breaking. Therefore, while the phase transition is accompanied by a change in the tetragonal axial ratio, we will consider only the distortion illustrated in fig 1 and keep the lattice constant fixed at the cubic value which we choose to be 5.18Å. We are able to compare our results with *ab initio* density-functional calculations using a new, high precision, band-structure program in a basis of smooth Hankel functions augmented in atomic spheres, with plane wave expansions of the potential and charge density in interstitial regions [14]. For ZrO_2 we use nine $2s2p3d$ oxygen basis functions at kinetic energy $\kappa^2 = -0.2$Ry and a further s function at $\kappa^2 = -0.5$, augmented in spheres of radius 1.8 a.u. Zr basis functions are $5s$ at $\kappa^2 = -0.01, -0.5, -2$, $4p$ at $\kappa^2 = -0.01, -0.4, -2$ and $4d$ at $\kappa^2 = -0.1, -0.4$. The Zr $5p$ are too high in energy to contribute significantly to the energy bands and the Zr $4s$ dispersion is treated in a frozen, overlapping core approximation. The radius of the Zr atomic spheres is 2.1 a.u.
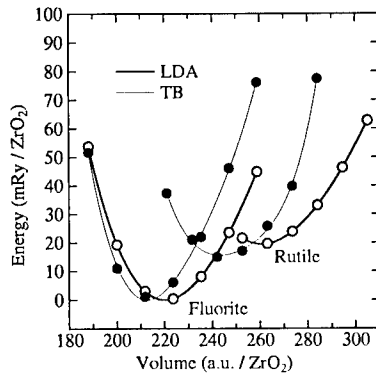
FIG 3: Energy volume curves for fluorite and rutile ZrO_2 calculated in the local density approximation (open circles) and in the new tight-binding model (filled circles). All electronic structure parameters were adjusted to reproduce the energy bands (fig 4) and so only the two parameters in the Born-Mayer repulsive potential were available to fit to the minima and curvature of the ab initio data.

We have attempted to make our tight-binding model for ZrO_2 as simple as possible. Therefore we employ a minimal basis of only $4d$ orbitals on Zr and $2p$ and $3s$ on oxygen. In this way we are describing only the $2p$ and not the $2s$ valence bands of oxygen and using the O–$3s$ states to describe the higher lying bands and also to mix with the $2p$ to provide the polarisability of the oxygen ion. Relative to the energy of the O–$2p$ level, the O–$3s$ atomic

270

level is at 0.35Ry, and the Zr–4d at 0.1Ry. We use a value of $U = 1$Ry on all orbitals. We allow only nearest neighbour matrix elements between oxygen atoms, namely $ss\sigma$, $sp\sigma$, $pp\sigma$, $pp\pi$ and have used length scaling adjusted to reproduce the volume dependence of the O–2p band width in an *ab initio* calculation. The only other inter-site matrix elements are those connecting Zr and oxygen, namely $sd\sigma$, $pd\sigma$ and $pd\pi$. These scale in the usual canonical way [11]. Only in the case of the rutile structure, shown in fig 2 do we also allow Zr-d–Zr-d matrix elements (between the two-fold coordinated Zr ions along the c-axis) and these are also chosen canonically. The magnitudes of the hopping matrix elements are chosen to approximately reproduce the *ab initio* band widths. The pair potential we employ is of the Born–Mayer form: $\phi(r) = Ae^{-\lambda r}$, between oxygen and Zr nearest neighbours only, and we have chosen A and λ to reproduce approximately the *ab initio* energy volume curves of ZrO$_2$ in the fluorite and rutile structures. These are shown in fig 3. Comparison between the tight-binding and *ab initio* energy bands are shown in fig 4. We can discuss our choices of the $\Delta_{\ell''\ell'\ell}$ parameters in terms of these.

First we mention those pertaining to the Zr sites. The most important is Δ_{ddg} and this is the integral denoted $\langle r^4 \rangle$ in crystal field theory of transition metal ions [10]. In an environment with cubic point symmetry, the first non vanishing component of the potential in the expansion (3) is for $\ell = 4$ and it is this which breaks the degeneracy of the five d-orbitals and splits them into the t_2 and e manifolds. In an octahedral or tetrahedral environment the t_2 states are split below the e and *vice versa* in a cubic field. Our approach incorporates this feature very naturally for a single value of Δ_{ddg} as seen in fig 4, where in fluorite the t_2 bands are higher and in rutile they are lower than the e bands. A value of $\Delta_{ddg} = 65$ a.u. achieves proper agreement between tight-binding and *ab initio* bands. The only other Δ-parameter needed is Δ_{ddd} and we find the bands quite insensitive to its choice. We set $\Delta_{ddd} \approx \sqrt{\Delta_{ddg}}$.

On the oxygen sites, in the minimal basis there are also two Δ-parameters. These are Δ_{spp} and Δ_{ppd}. In fluorite, there are no dipole or quadrupole moments on the oxygen ions and these integrals do not contribute to the bandstructure. The O-2p bandwidth is controlled by the magnitude of the hopping integrals. However rutile is necessarily distorted in the sense that the space group symmetry is tetragonal, while the oxygen octahedra are imperfect. Therefore there are both dipole and quadrupole moments on the oxygen sites. Whereas the bands are found to be insensitive to the value of Δ_{spp}, we find that it is Δ_{ppd} that controls the width of the O-2p band due to the splitting of the O-p levels by the field gradients. This exposes an essential difference between the fluorite and rutile modifications of ZrO$_2$ which emphasises the value of our model used in combination with *ab initio* band calculations. We will defer attaching a value to Δ_{spp} until the next paragraph.

The next and final question we will address is the role of the electronic structure in the spontaneous distortion of the fluorite structure to the tetragonal phase—modelled in the present work as the change in total energy along the one dimensional path as the oxygen atomic columns move alternately up and down as shown in fig 1, while keeping the lattice constants fixed. The total energy calculated *ab initio* is shown in fig 5, and this calculation predicts the same minimum as an earlier density-functional calculation using full potential LAPW [15]. Since the dipole moment arising from the symmetry breaking contributes to the amount of force on the oxygen ions, we now adjust the final Δ-parameter, Δ_{spp} on the oxygen sites to obtain as close agreement as we can with the *ab initio* energy. We see that the model correctly predicts the symmetry breaking phase transition from fluorite to the tetragonal modification.
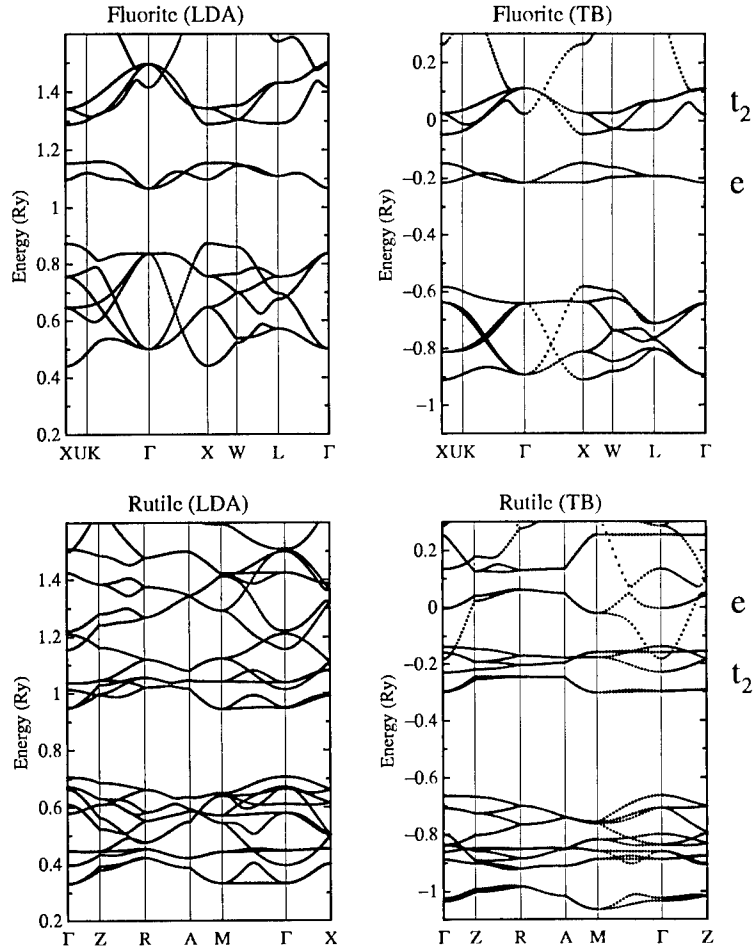
FIG 4: *Energy bands in fluorite and rutile ZrO_2 using local density approximation (left) and the new tight-binding model (right). The energy zeros are arbitrary. In all four figures, the lowest lying set of bands are the occupied oxygen 2p valence bands. Unoccupied Zr 4d and oxygen 3s bands are separated from these by a band gap. Features to note are (i) the good reproduction of the fluorite LDA bands in the minimal basis tight-binding approximation. The free-electron-like parabola at Γ above and hybridised with the t_2 manifold is the oxygen 3s band; (ii) again, in fluorite, the cubic symmetry demands that the only relevant Δ-parameters are Δ_{ddd} and Δ_{ddg} on Zr. The former has little effect on the bands but it is Δ_{ddg} that produces the large crystal field splitting of the 4d bands; (iii) in rutile the sign of the cubic splitting is reversed, and this effect is reproduced naturally in our structure constant expansion without explicit fitting; (iv) whereas the width of the 2p bands in fluorite is determined by the size of the pp and pd hopping integrals, in rutile the width is largely determined by the size of Δ_{ppd} which we adjust to reproduce the LDA bandwidth.*
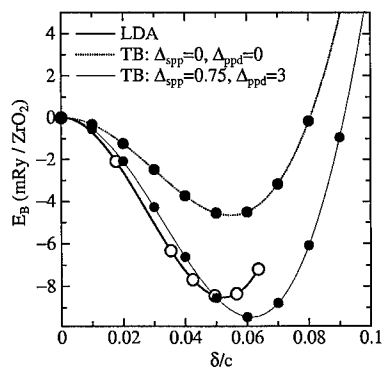
FIG 5: Energy as a function of the displacement δ of oxygen atoms in units of the tetragonal c lattice parameter (in our case equal to the fluorite lattice constant)—see fig 1. Open circles are the ab initio *LDA data. We show two sets of data from the new tight-binding model. The thin full line are data employing Δ-parameters (in a.u.) as shown. Δ_{ppd} has already been fixed to obtain the correct width of the valence band in rutile. Δ_{spp} which produces the dipole moment on the oxygen atom has been adjusted to get agreement with the LDA data here. The broken line shows the energy in the point charge self-consistent tight-binding approximation, where the Δ-parameters are set to zero. The distortion persists in this limit although the energy gain is less. This demonstrates that the distortion is not a consequence of the polarisablility of the oxygen ions alone, but that this does play a major rôle.*

DISCUSSION AND CONCLUSIONS

As far as we know, this is the first time that tight-binding theory has been extended to include polarisability and crystal field splitting by direct calculation of the charge and potential expanded in angular momentum components. With only a small number of physically motivated parameters, one may now include these effects in order to address topics such as those mentioned in the introduction. We expect there to be many more potential applications.

In the example we have chosen, we have shown firstly how a very small set of basis functions, properly chosen, is adequate to obtain the bandstructure of fluorite ZrO_2 in agreement with *ab initio* bands. By including the parameterised crystal field Δ-integrals, the Zr–$4d$ bands split naturally into t_2 and e sub bands, and this splitting is reversed in rutile simply as a consequence of using a structure constant expansion for the potential, without the need of explicit parameterisation. Futhermore in rutile we see the importance of crystal field effects upon the width of the O–$2p$ band, a fact that would be hard to deduce without working within a model such as ours to interpret *ab initio* bands. By choosing just two parameters in a simple pair potential, our model is able to reproduce the fact that the rutile structure is less stable than fluorite; consistent with rutile never being observed as a pseudomorph of ZrO_2 despite this being the stable phase of TiO_2. This competition between rutile and fluorite is problematic in many atomistic descriptions in terms of purely classical electrostatic terms, and this is a consequence of the ionic radius of Zr^{4+} placing ZrO_2 on the border between prefering six or eight-fold coordination by oxygen if one uses the usual radius ratio rules [7].

273

Finally the model correctly predicts the spontaneous symmetry breaking in the fluorite structure.

We can also use the results of our calculations to address the question of the *origin* of the symmetry breaking. In a non quantum mechanical description [7], the symmetry breaking introduces extra degrees of freedom, particularly the quadrupole polarisation of the oxygen ion and this allows the crystal to lower its energy. We can test this hypotheses by calculating the energy path after setting the Δ-parameters to zero (*ie*, returning to the point charge self-consistent tight-binding approximation). As seen in fig 5, the symmetry breaking persists in this limit indicating bandstructure effects at work in addition to classical electrostatic forces. However, the effect is weakened indicating the role of the ionic polarisability in the phase transition.

We expect this preliminary investigation to lead to further applications of our new formalism. Including polarisation of atomic charges in the tight-binding approximation will allow calculations of lattice dynamics and defect structure in oxides and other ceramic materials. We anticipate also applications to surface structure and defects as well as work functions in metals. A combined approach to metals and ceramics will assist in our understanding of the metal ceramic interface and the nature of bonding in such heterogeneous solid state systems. New applications in molecular physics may also be realised.

REFERENCES

1. R. C. Kittler and L. M. Falicov, *Phys. Rev. B*, **18**, 2506 (1978); M. O. Robbins and L. M. Falicov, *Phys. Rev. B*, **29**, 1333 (1984).
2. W. A. Harrison, *Phys. Rev. B*, **31**, 2121 (1985).
3. J. A. Majewski and P. Vogl, *Phys. Rev. Lett.*, **57**, 1366 (1986); *Phys. Rev. B*, **35**, 9679 (1987).
4. A. T. Paxton, *J. Phys. D*, **29**, 1689 (1996)
5. C. Priester, G. Allan and M. Lannoo, *Phys. Rev. B*, **33**, 7386 (1986).
6. A. P. Sutton, M. W. Finnis, D. G. Pettifor and Y. Ohta, *J. Phys. C*, **21**, 35 (1988)
7. M. Wilson, U. Schönberger and M. W. Finnis, *Phys. Rev. B*, **54**, 9147 (1996).
8. D. J. Chadi, in *Atomistic simulation of materials*, edited by V. Vitek and D. J. Srolovitz (Plenum Press, New York, 1989), p. 309.
9. A. J. Stone, *The theory of intermolecular forces*, (Oxford University Press, 1996).
10. J. S. Griffith, *Theory of transition metal ions*, (Cambridge University Press, 1960); D. S. McLure, in *Phonons*, edited by R. W. H. Stevenson (Oliver and Boyd, London, 1966); A. M. Stoneham, *Theory of defects in solids*, (Oxford University Press, 1975).
11. O. K. Andersen, in *NATO ASI: The electronic structure of complex systems*, edited by P. Phariseau and W. M. Temmerman, (Plenum, New York, 1984), p. 11.
12. M. Methfessel, *Multipole Green functions for electronic structure calculations*, (Katholieke Universiteit te Nijmegen, 1986).
13. D. J. Chadi, *Phys. Rev. B*, **41**, 1062 (1978)
14. M. Methfessel and M. van Schilfgaarde, *NFP Manual 1.00*, (IHP-Frankfurt/Oder, 1997).
15. H. J. F. Jansen, *Phys. Rev. B*, **43**, 7267 (1991).

HOW FAR TO USE TIGHT-BINDING POTENTIALS
FOR BIMETALLIC SURFACE MODELLING ?

G. TRÉGLIA*, AND B. LEGRAND**
* C.R.M.C.2 - C.N.R.S., Campus de Luminy, Case 913, 13288 Marseille Cedex 9, France
** SRMP/DECM, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France

ABSTRACT

Modelling in a realistic way both equilibrium and dynamical processes on bimetallic surfaces requires the availability of interatomic potentials sufficiently simple (i.e. analytical) although derived from the electronic structure. This is possible in the framework of Tight-Binding formalism. We present here a review of the applications of such potentials, together with some reflexions about their limitations.

INTRODUCTION

The aim of this lecture is to give tools allowing to model surface processes in bimetallic systems from their electronic structure. They include both atomic (reconstructions, relaxations, superstructures) and chemical (segregation, incorporation, interdiffusion) rearrangements, occuring not only **at** but also **out of** equilibrium. This requires to couple analytical models (Mean Field) and numerical simulations (Molecular Dynamics, Monte Carlo) with interatomic potentials both simple (i.e. analytical) and realistic (which means grounded on electronic structure). This is possible within Tight-Binding (TB) formalism, provided one can develop simultaneously *ab initio* approaches giving "exact" references to validate (or modify) the underlying assumptions of semi-empirical models, which is the most crucial as the complexity of the system increases (bimetallic surfaces).

In the following, we will first illustrate the ability of many-body TB-potentials, derived in the Second Moment Approximation (TBSMA), to model the atomic structure of pure metal sytems with low dimentionality (surface relaxation and reconstruction, cluster shapes). Then, we will show how a second type of potentials, Ising-like (TBIM: Tight-Binding Ising Model), can be derived to describe the chemical structure of alloy surfaces (competition between surface segregation and bulk tendency to ordering or phase separation) at thermodynamical equilibrium. We will then extend these two approaches to study dynamical processes involved during the growth (homo and hetero epitaxy), such as diffusion (near steps and on terraces), incorporation, interdiffusion and *surface alloy* formation. Finally, we will show the limits of these approaches, both from the point of view of their consistency and from the comparison with *ab initio* calculations.

TB-POTENTIALS FOR SURFACE ATOMIC STRUCTURE

Tight-binding Second Moment Approximation (TBSMA)

Let us start from the usual tight-binding Hamiltonian which, in the basis of atomic orbitals λ at site n $|n,\lambda\rangle$, can be written [1]:

$$H = \sum_{n,\lambda}|n,\lambda\rangle(\varepsilon_{at}^{\lambda} + \alpha^{\lambda})\langle n,\lambda| + \sum_{n,m,\lambda,\mu}|n,\lambda\rangle\beta_{nm}^{\lambda\mu}\langle m,\mu| \tag{1}$$

in which $\varepsilon_{at}^{\lambda}$, α^{λ} and $\beta_{nm}^{\lambda\mu}$ are respectively the atomic level, crystal field and hopping integrals, the latter being rapidly damped (after 1st or 2nd neighbours). An essential advantage of this formalism is to define in a simple way the local density of states at a given site n_0,

$$n_{n_0}(E) = \lim_{\eta \to 0^+} \left[-\frac{Im}{\pi} \sum_{\lambda} \langle n_0, \lambda | G(E + i\eta) | n_0, \lambda \rangle \right] \qquad (2)$$

without resorting to any periodicity condition (n_0 can be a defect site), from the projection of the Green function: $G(z) = (z-H)^{-1}$. $G(z)$ writes as a continued fraction [2], the coefficients of which are derived either from the knowledge of the p first moments μ_p of $n_{n_0}(E)$:

$$\mu_p(n_0) = \int_{-\infty}^{+\infty} E^p n_{n_0}(E) dE = \sum_{\lambda} \langle n_0, \lambda | H^p | n_0, \lambda \rangle \qquad (3)$$

obtained by counting paths on the lattice[2], or directly by constructing a new basis tridiagonalising H within the so-called *recursion* method [3]. The continued fraction is then terminated by the asymptotic values of the coefficients, which are related to band (and/or gap) edges and fitted to the band structure [4]. Obviously, $n_p(E)$ is the most precise as the number of exact coefficients is large. It is worth noticing that for elements at the end of the transition series, it is necessary to take into account the s and p valence electrons and their hybridization with the d ones ($\lambda = 1,9$) to get a density of states in good agreement with that derived from ab initio calculations [5]. Then, the cohesive energy of the system writes [6]:

$$E_{coh} = \sum_{n,\lambda} \int^{E_F} (E - \varepsilon_{at}^{\lambda}) n_n^{\lambda}(E) dE + \frac{1}{2} \sum_{n \neq m} \iint dr dr' \frac{Q_n(r) Q_m(r')}{|r - r'|} \qquad (4)$$

where $n_n(E)$ and $N_n \left(= \int^{E_F} n_n(E) dE \right)$ are respectively the local density and the charge at site n. Assuming charge neutrality implies that $N_n = Z_n$ (ionic charge). The first term in the r.h.s. member is the band energy (E_{band}) and the second one the pair interaction (E_{rep}) between neutral atoms with charge density: $Q_n(r) = Z_n \delta(r-n) - N_n(r-n)$. Unfortunately, E_{rep} is not sufficient to account for the repulsive part of the energy. Actually, the tight-binding approximation fails to reproduce repulsions at short distances since it does not account for the non orthogonality of wave functions on different sites and for the compression of sp electrons which play an important role before that the Coulomb repulsion (1/R) becomes really efficient. To go beyond this dificulty, the idea is to build a semi-phenomenological tight-binding model in which the band part, coming from the electronic structure, has a many-body character whereas the repulsive one is a pairwise potential fitted to some physical properties. The equation (4) then reduces to:

$$E_{coh} = E_{band} + E_{rep} \qquad (5)$$

Since the integral does not depend on details of n(E), the band term can be calculated from a schematic rectangular density of states under the single assumption that it has the same second moment (TBSMA) as the exact one, which leads to [7]:

$$E_{coh} = -\beta \sqrt{\sum_R e^{-2q(\frac{R}{R_0}-1)}} + A \sum_R e^{-p(\frac{R}{R_0}-1)} \tag{6}$$

where β is an "effective" hopping integral, corresponding to the 1st neighbour distance R_0. In practice, we will consider the four parameters on the same footing, and determine them, for a given metal, to fit experimental quantities such as the cohesive energy, lattice parameter and elastic constants or the so-called *universal equation* [8]. Note that this TBSMA-potential is very similar to those derived elsewhere within the Embedded Atom Model (EAM [9]) or within the Glue Model [10]. Its main advantage compared to the latter is its physical transparency which clearly shows its limitations... and then its possible improvements (increasing the number of exact moments). However some features are wrong. Thus, the shear elastic constants depend significantly on the number of moments used. In the TBSMA, it is found too weak for bcc metals with Ne=4 (V, Nb et Ta) and too large for fcc ones around Ne=8 [6].

Application to surfaces:

Due to the broken bonds, the surface atoms can undergo displacements with respect to their bulk positions. In all cases, there is at least a vertical relaxation, which is experimentally known to be *inwards* (contraction of the first interlayer distance) for transition metals. This is a first success of TBSMA-potential to reproduce this contraction, which is found of a few percents and proportional to the number of broken bonds, contrary to simple pair potential models which predict an *outwards* relaxation [11]. This comes from the stronger decrease with coordination Z of the repulsive term compared to the attractive one ($\sim\sqrt{Z}$). One has then to use such many body potentials if one aims to study surface atomic rearrangements.

In some cases, not only vertical but also lateral atomic rearrangements can occur, changing the two-dimensional periodicity and leading to so-called surface reconstructions. Unfortunately, it is less easy to analyze in a simple way the possible lateral atomic rearrangements since they involve both increasing and decreasing distances. Nevertheless, one can get some trends which can be either along the transition series (zig-zag reconstruction of the (100) face of bcc crystals occuring for column VIa but not for Va) or along a column (missing row reconstruction of fcc (110) ones which is observed for the 5d series only). Both trends are well interpreted in the framework of tight-binding calculations. However, although the missing row reconstruction of the (110) surface is well reproduced by Quenched Molecular Dynamics within TBSMA potentials [12], due to the element dependence of the q parameter which drives the distance dependence of the hopping integrals ($q_{Cu} < q_{Ag} < q_{Au}$), understanding the zig-zag reconstruction requires to go beyond second moment arguments. Indeed, it is due to the broadening of the quasi-atomic surface peak of the local bcc (100) density of states under the lattice distortion, which leads to an energy gain for d band filling around 5 (middle of the series) [13].

Application to clusters

Due to the tremendous interest of clusters of late transition and noble metal elements for the reactivity rates in heterogeneous catalysis, a lot of experimental and theoretical works have been performed on these systems, in particular concerning their atomic structure. All the theoretical studies using geometrical optimisation procedures to relax the structure find that the icosahedron is more stable than the structures based on the fcc lattice (cuboctahedron and Wulff polyhedron) in a more or less large size range. From the experimental point of view, the critical size varies between some tens (Au) and some thousands (Ag). These results are consistent with Quenched Molecular Dynamics calculations performed within the TBSMA model since Ag clusters are found to have their energy minimum with the icosahedral shape until 309 atoms size while Au clusters are stabilized with the fcc Wulff polyhedron morphology for all sizes apart from the 13 atoms size for which the icosahedral structure is preferred relatively to the cuboctahedral one [14].

Moreover, TBSMA calculations show that this critical size strongly depends on the local relaxation, especially concerning the icosahedron. Indeed the relaxation profile differs remarkably between fcc-type and icosahedral clusters which are found to adopt a very inhomogeneous atomic relaxation. The first ones are characterized by a contraction of the outer intershell distance only, whereas this contraction is present both for the outer and the inner shells in icosahedron. More precisely, a considerable core contraction is found in the icosahedron case, which increases as a function of the cluster size, whereas it is quite negligible in the cuboctahedron case [14].

A spectacular consequence of this contraction of the inner shells for the icosahedral structure is the existence of a strong compressive pressure in the core which can be relaxed by pulling out one or more atom(s), i.e. by introducing constitutive vacancies. Indeed, for noble metals, one finds that the formation energy of a single vacancy at the center of an icosahedral structure decreases as the size increases, up to becoming negative(for $N \geq 309$ atoms in Cu), which means that the cluster with a central vacancy is more stable than the full one. At the opposite, such an effect is not found for fcc-type clusters (cuboctahedron as Wulff polyhedron) for which it tends monotonously to its value in the bulk. Since the stability of the vacancy increases with the cluster size, one can wonder to what extent a cavity of many atoms could be stabilized for larger sizes. In fact, a fourfold tetrahedral shaped cavity becomes even more stable than the single one when the size increases, but beyond the morphological transition to fcc-type structures. Finally, let us mention that one finds a tendency along the noble metal column: the vacancy formation energy becomes negative more and more early (that means towards the small sizes) from Cu to Au for one vacancy as for a larger cavity. Taking into account the stability domain of the icosahedron relatively to fcc structure (namely the Wulff polyhedron), one concludes that there is a stability range of size for Cu and Ag icosahedra with a central constitutional vacancy but not for Au ones [14].

TB-POTENTIALS FOR BIMETALLIC ALLOY SURFACES

Tight-Binding Ising Model (TBIM)

For a bimetallic system A_cB_{1-c}, the tight-binding hamiltonian depends on the chemical configuration $\left\{p_n^i\right\}$, where $p_n^i = 1$ if site n is occupied by atom i (= A, B) and $p_n^i = 0$ if not. If one neglects any possible "off-diagonal disorder" due to the difference in band width or atomic radius between the two constituents (i.e. the ij-dependence of the hopping integrals),

with respect to the (most important) "diagonal disorder" coming from the variation of the atomic level ε^i as a function of the element, it writes [6]:

$$H = \sum_i \sum_{n,\lambda} |n,\lambda\rangle p_n^i \varepsilon_i^\lambda \langle n,\lambda| + \sum_{n,m,\lambda,\mu} |n,\lambda\rangle \beta_{nm}^{\lambda\mu} \langle m,\mu| \qquad (7)$$

Introducing a surface is now more complicated since not only atomic but also chemical rearrangements can occur. Unfortunately, accounting for "diagonal disorder" requires to go beyond second moment (at least four) [6]. It is then necessary to develop an other type of approximation, which will only be suited to study the effect of chemical ordering on a rigid lattice. As for pure metals the total energy of the alloy, for a given configuration, cannot be described as a sum of pair interactions. Nevertheless, the (small) part of the energy which depends explicitly on configuration (essential in ordering problems) can be written in an Ising-like form by developing the energy in a pertubative way [15] with respect to that of the disordered state ($\overline{E}(c)$), described within the Coherent Potential Approximation (CPA [16]). In presence of a surface, assuming a local charge neutrality condition, one obtains the so-called Tight-Binding Ising Model (TBIM) [17]:

$$E_{coh}\left(\left\{p_n^i\right\}\right) = \overline{E}(c) + H^{mix}\left(\left\{p_n^i\right\}\right) \qquad (8)$$

$$H^{mix}\left(\left\{p_n^i\right\}\right) = \sum_{n,i} p_n^i h_n^i + \frac{1}{2} \sum_{n,m,i,j} p_n^i p_m^j V_{nm}^{ij} \qquad (9)$$

in which a linear term and a quadratic one appear:

$$h_n^i = \frac{Im}{\pi} \int^{E_F} dE \sum_\lambda \log\left[1 - \left(\varepsilon^i - \sigma_n(E)\right)\overline{G}_{nn}^{\lambda\lambda}(E)\right] \qquad (10)$$

$$V_{nm}^{ij} = -\frac{Im}{\pi} \int^{E_F} dE t_n^i(E) t_m^j(E) \sum_{\lambda\mu} \overline{G}_{nm}^{\lambda\mu}(E)\overline{G}_{mn}^{\mu\lambda}(E) \qquad (11)$$

in which $\overline{G}_{nm}^{\lambda\mu}(E) = \langle n\lambda|\overline{G}(E)|m\mu\rangle$ are interatomic average Green functions, calculated from the electronic structure of the disordered state in the CPA approximation, and $\sigma(z)$ the effective potential (replacing ε^i) determined from the self-consistency CPA condition [16]:

$$\sum_i c^i t_n^i(E) = 0 \qquad \text{with} \qquad t_n^i(E) = \frac{\varepsilon^i - \sigma(E)}{1 - \left(\varepsilon^i - \sigma(E)\right)\sum_\lambda \overline{G}_{nn}^{\lambda\lambda}(E)} \qquad (12)$$

Due to broken bonds, the equilibrium concentration at the surface has no reason to be the same as in the bulk, which leads to the phenomenon of surface segregation [18]. For an $A_c B_{1-c}$ binary alloy, a single configuration parameter $p_n = p_n^A$ is required since: $p_n^A + p_n^B = 1$. The concentration profile $\{c_p\}$, where $c_p = <p_n>$ (for any site n in the pth plane parallel to the surface: p=0), is assumed to be homogeneous and determined as the one which minimises the free energy, for instance in the mean-field approximation. Let us define the chemical potential

279

as $\mu = \mu^A - \mu^B$ and denote N_p the number of atoms in the p-plane, each one having Z_p ($=Z^{tot}$ in the bulk) first neighbours among which Z_{pq} in the q-plane. One can then write:

$$F = \langle H^{mix} \rangle - TS - \sum_p N_p (c_p - c)\mu \tag{13}$$

$$\forall c_p, \quad \frac{\partial F}{\partial c_p} = 0 \quad ====> \quad \frac{c_p}{1 - c_p} = \frac{c}{1 - c} \exp\left(-\frac{\Delta E_p^{TBIM}}{kT} \right) \tag{14}$$

where the segregation energy, which accounts for the exchange of a A bulk atom with a B surface one, is defined as:

$$\Delta E_p^{TBIM} = \Delta h_p + (1 - 2c)Z^{tot}V - \sum_{p'=-q-p}^{p'=+q-p} (1 - 2c_{p+p'})Z_{p,p+p'}V_{p,p+p'}$$

with: $\tag{15}$

$$\Delta h_p = (h_p^A - h_p^B) - (h_{bulk}^A - h_{bulk}^B) \quad \text{and} \quad V_{nm} = \frac{\left(V_{nm}^{AA} + V_{nm}^{BB} - 2V_{nm}^{AB} \right)}{2}$$

V_{pq} is the value of the "effective" alloy pair interaction V_{nm} for sites n and m belonging to planes p and q respectively. It is worth noticing that V_{nm} can be used to calculate the mixing or ordering energies of the system under study, but **not** its cohesive energy ! Its sign indicates the tendency to order (V>0) or phase-separate (V<0). V_{nm} decreases rapidly with the distance (n-m) ($V_1 \gg V_2$, V_3, $V_4 \gg V_5$, ... for fcc structure [19]). Moreover it depends on bulk concentration (which could change the tendency to order or phase separate in a system) and on the average d band filling $\overline{N}_e = cN_e^A + (1 - c)N_e^B$. In fact, alloys with a nearly half-filled band tend to order whereas those with a nearly filled or empty bands tend to phase separate [4,6,15,19]. At the surface, these effective pair interactions are enhanced with respect to the bulk, $1.5V < V_1 < 2V$ [17].

In what concerns the local term Δh_p, it vanishes in the bulk (p>>0), except for complex unit cells for which, nevertheless, it is quite negligible [20]. At the surface (p=0), it is almost identical to the difference in surface energies ($\tau^A - \tau^B$) [17,21]. In fact, Δh_0 is the main driving force which leads to the segregation of the element with the lowest surface energy. Finally, let us note that, up to now, the derivation of TBIM has been performed on a rigid lattice, which is probably too crude in the case of large size mismatch between the constituents. However, it is possible to introduce this effect by adding a third contribution to the segregation energy:

$$\Delta E_p = \Delta E_p^{TBIM} + \Delta E_p^{size}(c) \tag{16}$$

with $\Delta E_p^{size}(c) \neq 0$ if p=0 (and 1 for open surfaces). $\Delta E_0^{size}(c)$ is calculated in both dilute limits (c→0,1) in the framework of TBSMA, by determining the four mixed A-B parameters in order that A and B only differ by their size [21,22]. This leads to a contribution which significantly differs from the one derived from elasticity theory since the latter leads in both limits to the segregation of the impurity, whatever its size. On the contrary, the tight-binding

term is found strongly asymmetric, leading to a segregation of the impurity when it is the largest only (at least for close-packed surfaces). This comes from the anharmonicity of the potential which exhibits a strong asymmetry between tensile and compressive pressures.

Equilibrium segregation at alloy surfaces

Determining the concentration profile then requires to solve the non linear coupled equations (14), in which ΔE_p^{TBIM} has to be replace by the complete segregation energy ΔE_p (16). An essential problem comes from the non linearity which allows many solutions to exist, so that one has not only to be sure to get the one with the lowest energy (thermodynamic equilibrium) but also to find the metastable ones which could be stabilized under variation of external parameters (pressure, temperature), leading to phase transitions. This can be done in the framework of *phase portrait* or *local field* methods [21].

At high temperature, the segregation leads to an enrichment of the surface plane by the element for which the surface driving forces (surface energy and size mismatch) are the most favourable, the nature of the profile being determined by the sign of the ordering bulk term (V). Thus, for ordering systems (V>0) the first underlayer is depleted with respect to the segregating species (oscillating profile), whereas it is still enriched but to a lesser extent (monotonous profile) for phase separating ones. Then the profile is damped on a few layer up to bulk concentration [21]. At low temperature for V<0 (below the critical temperature of phase separation), surface segregation is replaced by preferential surface precipitation. For V>0, if the ordered phase can be analysed as an alternate stacking of inequivalent planes parallel to the surface (e.g. .../A/B/A/B/ ... for c=0.5), it determines the nature of the most favourable termination ... the other possible terminations being found metastable.

The most interesting situation is found for intermediate temperatures, namely slightly above bulk critical temperatures. Thus, for V<0 with a strong tendency to solute segregation (e.g. $Cu_{1-c}Ag_c$ for c→0), layering transitions occur in which the initially almost pure Cu surface layers become successively almost Ag pure when approaching the solubility limit, leading to a complete wetting of the surface, i.e. a "pre-phase separation" [23]. Note that the first layering transition has indeed been observed experimentally [24].

For ordering systems, the situation differs according to the respective contributions of surface and bulk driving forces. If the former are not too high (PtNi, for which $\Delta h_0 \sim 0$), the competition between the asymmetric size effect and the ordering one leads to "profile phase transitions" between pure Ni or Pt terminated surface sandwiches, as a function of surface orientation, concentration and temperature [25], which have indeed been experimentally observed [18]. If they are predominent, one observes blocking of surface concentration on some "magic" values (c_0=1/4, 1/3, 1/2, ...), characteristic of two-dimensional ordering [26].

Surface superstructure

Let us recall that in the above description, the atomic relaxation has only be taken into account by means of the dilute size-effect in segregation energy. However, for strong size-mismatch and surface enrichments, one could expect some important atomic rearrangements. Thus, for V<0 and a strong segregation of the A element at the surface of a dilute B(A) alloy, one can expect reconstructions of the surface plane when the size-mismatch is large. The corresponding superstructures can be found by using Molecular Dynamics simulations with TBSMA potentials, in which the four mixed A-B parameters are now derived from alloy thermodynamic quantities such as mixing or solution energies. This means that, being unable

to account properly for diagonal chemical order, we simulate it by an effective off-diagonal one. Thus, one finds for Cu(Ag) (111) a large surface undulation [27], which is confirmed by recent STM experiments [28].

For V>0, another competition between chemical and geometrical order can occur in presence of steps at the surface. Indeed, for vicinal surfaces with single steps, an alternate stacking of pure planes, .../A/B/A/B/..., in the ordered bulk will present simultaneously stable and metastable terminations at the surface. Then, depending on the relative forces of bulk and surface forces, the surface could present either single steps and some unfavourable terminations or double steps and then a single favourable termination ... the price to pay being to break more bonds! This problem does not arise at high temperature since the disordered alloy has not to preserve the stacking. Indeed, a transition from double to single steps, associated to the order-disorder transition, has been experimentally evidenced in CuPd [29].

Bimetallic clusters

The coupling between segregation and reconstruction should be particularly important in bimetallic clusters. First, for small sizes, a finite matter effect occurs which means that, for dilute systems, the available quantity of segregant matter could be lower than the quantity of surface sites. In addition, geometrical frustrations can appear, due to the coexistence of facets with different orientations. This imposes to mix concentration profiles which could be not compatible (due to ordering tendency), leading to antiphases boundaries. As a result the local order at the surface could be different from that of semi-infinite systems. Finally, one expects the segregating species to enrich progressively the low coordination sites (vertices, edges, ...).

In practice, coupling between chemical and atomic structure is now too strong to allow us to separate them. Therefore, the energetic model has to combine both TBIM (segregation and ordering) and TBSMA (distance dependence of the interatomic potential) in order to be able to treat various morphologies (fcc polyhedra, bcc dodecahedra, icosahedra) and to study the competition or synergy between bulk ordering and surface segregation. Thus, in the CuPd system [30], besides the confirmation of the above statements, the coupling between atomic relaxations and chemical arrangements in the icosahedral structure has been evidenced. In particular, for small sizes, if the usual sequence of relative stabilities (icosahedron, fcc, and -well above- bcc dodecahedron) are recovered in the disordered state, chemical ordering at low temperature drives a spectacular reversal in which the bcc structure is stabilized with respect to fcc by chemistry, the icosahedron being destabilized by chemical order. Moerover, the order of the transition varies at the surface: starting from ordered particles at low temperature, one observes a surface induced disorder with respect to inner sites. The transition is less abrupt for the latter than for bulk sites when the size decreases, which means that a first order bulk transition can evolve towards second order in small cluster cores.

TB-POTENTIALS FOR GROWTH DYNAMICS AND DISSOLUTION KINETICS

Surface diffusion: influence of steps

Understanding growth mechanisms implies the knowledge of migration processes of adatoms on stepped surfaces and islands. Thus, a layer-by-layer growth is only possible if adatoms of an upper terrace can go down sufficiently fast not to be trapped by other adatoms, which depends on the relative heights of the barriers for terrace diffusion and step

descent (Schwöbel barrier). Moreover, the shape of two-dimensional islands during growth is driven by diffusion along its step edges since, for facets with different orientations, a slower diffusion along one type of edge will lead to an accumulation of atoms which will make it disappear. This is the case for fcc (111) surfaces for which two kinds of high-symmetry straight steps exist, labelled A (square microfacet on the step riser) and B (triangular microfacet). The spectacular experimental transitions with temperature between different island shapes and growth modes [31] can then be attributed to the hierarchy between the different barriers for diffusion of adatoms: on the terrace, along and across step edges.

In general, these barriers are obtained from static calculations in which the total energy of the system is minimized along some diffusion paths. Using Quenched Molecular Dynamics with TBSMA-potentials for Au/Au and Ag/Ag (111) one finds that the two possible descent mechanisms (exchange and hopping) indeed present barriers which are significantly enhanced with respect to the one for diffusion on the terrace (Schwöbel effect). It turns out that in gold the barriers are very similar for both steps, the exchange one being slightly lower than that for hopping. In silver, there is a large anisotropy between the steps: the descent at step B is much easier than at step A, because of the low barrier of the exchange process at step B [32].

In fact, one can wonder if such static barriers are relevant to interpret experiments at finite temperature. Indeed, it is possible to calculate dynamic barriers, using Molecular Dynamics now at high temperature and drawing the rates of the various events (exchange, hopping, repulsion by the step) in an Arrhenius plot. One finds that the dynamic barrier for diffusion on the terrace is unchanged relative to the static one, whereas the one for descending steps by the exchange mechanism is significantly lowered [32]. This dynamic lowering of the Schwöbel barrier definitively favors the exchange mechanism with respect to the hopping one.

Adatom incorporation

Steps not only play an important role during homoepitaxy but also in heteroepitaxy. In the latter case, a new process may occur, namely the incorporation of some deposited adatoms into the substrate, for instance during annealing. Identifying the elementary processes responsible of such an incorporation is essential since it is the first step of interdiffusion (see below). From STM experiments performed on Mn/Cu (100) system, it is clear that steps play a major role for incorporation, in relation with the preferential formation of vacancies at the kinks. This can be modelled by calculating the activation energies for the various processes relevant to the production of vacancies at steps on the Cu(100) surface within TBSMA and then the production rates for vacancies both analytically and by means of Monte Carlo simulations. One finds that Mn is indeed incorporated at step edges, due either to the step fluctuations or to exchange processes. Then the migration of incorporated Mn away from step is mediated by the vacancies produced at the steps, which is sufficient to reproduce the asymmetry of the spatial distribution of incorporated Mn between the upper and lower terrace observed by STM [33].

Surface alloy formation: the Kinetic Tight-Binding Model (KTBIM)

Modelling the interdiffusion phenomena during growth and annealing of a deposit requires to take into account simultaneously kinetic (diffusion) and thermodynamical (segregation) effects. This is possible within the KTBIM (Kinetic Tight-Binding Ising Model [34]), which is a simple one dimensional model, consistent with the (TBIM) equilibrium

state, i.e. the steady state concentration profile corresponds to the equilibrium profile given by (14). Assuming homogeneous concentrations per planes parallel to the surface, their time dependence $c_p(t)$ results from a detailed balance between the incoming and outcoming fluxes:

$$\frac{dc_p}{dt} = c_{p-1}(1-c_p)\Gamma_{p-1\to p} - c_p(1-c_{p-1})\Gamma_{p\to p-1}$$
$$+ c_{p+1}(1-c_p)\Gamma_{p+1\to p} - c_p(1-c_{p+1})\Gamma_{p\to p+1} \qquad (17)$$

the exchange frequencies $\Gamma_{p\to q}$ between the planes p and q being related to the segregation energies ΔE_p and ΔE_q, given by (16), through:

$$\Gamma_{p\leftrightarrow p+1} = Z_{p,p+1}\nu \exp-\left(\frac{Q \pm \dfrac{\left(\Delta E_{p+1} - \Delta E_p\right)}{2}}{kT}\right) \qquad (18)$$

in which the sign ($\pm$) corresponds to $\Gamma_{p\to p+1}$ (+) and $\Gamma_{p+1\to p}$ (-) respectively. ν is a characteristic vibration frequency and Q the bulk diffusion activation energy. Obviously, $Z_{p,q}=0$ for p<0 or q<0 (i.e. above the surface layer). Therefore, one takes into account the three segregation driving forces, which allows us to predict the different behaviours which can be observed if one deposits N A-layers onto a B-substrate ($c_p(t=0)=1$ for $p\leq N$ and $c_p(t=0)=0$ for p>N). They completely differ depending on the ordering bulk term (V sign), and on the surface terms (Δh_0 and ΔE_0^{size}) which can favour either A or B segregation [35]:

- Systems with tendency to phase separation (V<0):

When one deposits the element which tends to segregate (Ag/Ni ; Ag/Cu), it stays at the surface before dissolving into the substrate within an clearly non Fickian "layer-by-layer" mode [36]. On the contrary, if the substrate atoms tend to segregate (Ni/Ag, Cu/Ag, Fe/Cu), one observes, for a thin deposit, a "surfactant effect" leading to deposited clusters buried under one or two floating substrate layers. The time evolution of these buried clusters is governed by a competition between Ostwald rippening and dissolution into the substrate [37]. For thicker deposits, depending on segregation forces, one observes a transition from the "surfactant effect" to the "layer-by-layer" dissolution. All these behaviours can be understood in terms of local equilibrium linking dissolution kinetics to the equilibrium profiles in alloy thin films [35].

- Systems with tendency to chemical ordering (V>0):

The most spectacular effect is the formation of superficial compounds of type AB_n in the surface selvedge, under dissolution of A/B. If the deposited element is the one which tends to segregate, these so-called *surface alloys* correspond to AB_n ordered compounds of the bulk phase diagram. On the contrary, when the substrate atoms want to segregate, some new compounds appear which correspond to metastable terminations of bulk phases. The sequence of *surface alloys* which appear during dissolution as a function of time and film thickness is also characterized in terms of local equilibrium [35]. Experimentally, they are evidenced from kinetic blockings of dissolution in AES, LEED or photodiffraction [38].

TB-POTENTIALS: HOW FAR ?

Once granted the ability of TB-potentials, it is worth pointing out that one has to take some caution before using them. First of all, let us recall that derivation of the above TB-potentials requires drastic assumptions concerning charge neutrality near the defects (surface, impurities). For pure metal surfaces, it is now admitted that, at least for d electrons, the charge self consistency indeed reduces to a local neutrality condition which determines a modification of the level $\delta\varepsilon_0$. This d level shift is almost rigidly followed by the core levels, as experimentally confirmed by core level spectroscopy [39]. Unfortunately, for fcc elements at the end of the transition series, it is necessary to introduce sp-d hybridization. This is possible in the tight-binding framework, but the results strongly depend on the charge neutrality assumption. *Ab initio* (FP-LMTO) calculations in the surface selvedge indicate that charge neutrality has to be achieved, not only per inequivalent site but also per orbital, which means that almost no charge redistribution occurs between sp and d orbitals [5]. The situation at the surface of the alloy is still more complicated since charge transfers are involved not only between geometrical inequivalent sites but also between A and B species. At least in the bulk, it has ben shown that the main role of charge transfers on the effective pair interactions (11) is to reduce the chemical diagonal parameter, leading to the same results as without charge self-consistency [40]. Moreover, the ability of TBIM (which assumes charge neutrality) to reproduce experimental segregation trends, whereas not self-consistent calculations fail to do it, seems to confirm the underlying charge assumptions of the former.

Another drawback of TB-potentials is that they are better justified when atomic and chemical structure rearrangements can be separated. Concerning the rather *ad hoc* treatment of the size effect term in TBIM, comparison with calculations of the segregation energy, using either complete TBSMA or LMTO in the dilute limit, confirm that the approximation used is quite justified. The situation is less clear with respect to the overall consistency between TBIM and TBSMA, particularly in what concerns the effective off-diagonal disorder introduced in the latter to describe ordering problems. In particular, fitting the mixed parameters on solute solution energies does not insure that more local ordering properties are preserved. This is illustrated by calculating V (in the TBIM spirit) either from the energy of dissolution of a single impurity, or by moving two isolated impurities toxards first neighbour positions. This has been simulated alternatively by TBSMA and FP LMTO approaches, for two systems presenting a significant size mismatch and tendency to order (CuPd) and to phase separate (CuAg) respectively. For $V>0$ (CuPd), one indeed recovers the same effective V interaction as from solution energy. Moreover, performing the calculation for surface sites indeed leads to an enhancement of the pair interaction $V_0 \sim 1.5V$. On the contrary, for $V<0$, one finds a sign reversal both within TB and *ab initio* calculations. Even though relaxation around the impurity brings some improvement, it does not rule out the whole discrepancy.

REFERENCES

[1] Friedel J., Physics of Metals **1**, Cambridge University Press, 1978.
[2] Gaspard J.P. and Cyrot-Lackmann F., J. Phys. C **6**, 3077 (1973).
   Cyrot-Lackmann F. and Ducastelle F., Phys. Rev. Lett. **27**, 429 (1971).
[3] Haydock R., Heine V. and Kelly M.J., J. Phys. C **5**, 2845 (1972); **8**, 2591 (1975).
[4] Turchi P., Ducastelle F. and Tréglia G., J. Phys. C **15**, 2891 (1982).
[5] Sawaya S., Goniakowski J., Mottet C., Saúl S., Tréglia G., Phys. Rev. B **56**, 12161 (1997)
[6] Ducastelle F., Order and Phase Stability in Alloys, North-Holland, 1991.
[7] Rosato V., Guillopé M. and Legrand B., Philos. Mag. A **59**, 321 (1989).
[8] D. Spanjaard and M.C. Desjonquères, Phys. Rev. B **30**, 4822 (1984).
[9] Foiles S.M., Baskes M.I. and Daw M.S., Phys. Rev. B **33**, 7983 (1986).
[10] Garofalo M., Tosatti E. and Ercolessi F., Surf. Sci. **188**, 321 (1987).
[11] Desjonquères M.C. and Spanjaard D., Concepts in Surface Physics, Springer, 1995.
[12] Guillopé M. and Legrand B., Surf. Sci. **215**, 577 (1989).
[13] Legrand B., Tréglia G., Desjonquères M.C. and Spanjaard D., J. Phys. C **19**, 4463 (1986)
[14] Mottet C., Tréglia G. and Legrand B., Surf. Sci. Lett. **383**, L719 (1997).
[15] Ducastelle F. and Gautier F., J. Phys. F **6**, 2039 (1976).
[16] Velicky B., Kirkpatrick S. and Ehrenreich H., Phys. Rev. **175**, 747 (1968).
[17] Tréglia G., Legrand B. and Ducastelle F., Europys. Lett. **7**, 575 (1988).
[18] Gauthier Y., Baudoing R., Surface Segregation & Related Phenomena, Boca Raton, 1990
[19] Bieber A., Gautier F., Tréglia G. and Ducastelle F., Solid State Comm. **39**, 149 (1981).
[20] Turchi P., Tréglia G. and Ducastelle F., J. Phys. F **13**, 2543 (1983).
[21] Ducastelle F., Legrand B. and Tréglia G., Progr. Theor. Phys. Sup. **101**, 159 (1990).
[22] Tománek D., Aligia A.A. and Balseiro C.A., Phys. Rev. B **32**, 5051 (1985).
[23] Saúl A., Legrand B. and Tréglia G., Phys. Rev. B **50**, 1912 (1994).
[24] Eugène J., Aufray B. and Cabané F., Surf. Sci. **273**, 372 (1992).
[25] Legrand B., Tréglia G. and Ducastelle F., Phys. Rev. B **41**, 4422 (1990).
[26] Gallis C., thesis, Orsay, 1997.
[27] Mottet C., Tréglia G. and Legrand B., Phys. Rev. B **46**, 16018 (1992).
[28] Aufray B., Göthelid M., Gay J.M., Mottet C., Landemark E., Surf. Sci. Lett. *in press*
[29] Barbier L., Salanon B. and Loiseau A., Phys. Rev. B **50**, 4929 (1994).
[30] Mottet C., Tréglia G. and Legrand B., *to be published*
[31] Bott M., Michely T. and Comsa G., Surf. Sci. **272**, 161 (1992).
[32] Ferrando G., Tréglia G., Phys. Rev. B **50** 12104 (1994); Phys. Rev. Lett. **76** 2109 (1996)
[33] Ibach H., Giesen M., Flores T., Wuttig M. and Tréglia G., Surf. Sci. **364**, 453 (1996).
[34] Senhaji A, Tréglia G, Legrand B, Barrett N, Guillot C, Villette B, Surf Sci **274** 297 (1992)
[35] Legrand B., Saúl A. and Tréglia G., Mat. Sci. Forum **155-156**, 165 (1994).
   Tréglia G., Legrand B. and Saúl A., Il Vuoto **XXV**, 32 N4 (1996).
[36] Saúl A., Mat. Sci. Forum **155-156**, 233 (1994).
[37] Roussel J.M., Saúl A.,Tréglia G. and Legrand B., Phys. Rev. B **55**, 10931 (1997).
[38] Bardi U., Rep. Prog. Phys. **57**, 939 (1994).
[39] Spanjaard D, Guillot C, Desjonquères MC, Tréglia G, Lecante J, Surf Sci Rep **5** 1 (1985)
[40] Tréglia G., Ducastelle F. and Gautier F., J. Phys. F **8**, 1437 (1978).

TIGHT-BINDING CALCULATIONS OF COMPLEX DEFECTS IN SEMICON-DUCTORS: COMPARISON WITH AB INITIO RESULTS

M. KOHYAMA, N. ARAI* and S. TAKEDA*
Department of Material Physics, Osaka National Research Institute, AIST, 1-8-31, Midori-gaoka, Ikeda, Osaka 563, Japan, kohyama@onri.go.jp
* Department of Physics, Graduate School of Science, Osaka Univeristy, 1-16 Machikane-yama, Toyonaka, Osaka 560, Japan

ABSTRACT

Complex defects in Si and SiC such as coincidence tilt boundaries, planar defects and self-interstitial clusters were dealt with by using the transferable tight-binding method for Si and the self-consistent tight-binding method for SiC. These results have been compared with *ab initio* calculations of similar configurations. Essential features of the tight-binding results have been supported by the *ab initio* results. Especially, the agreement on stable atomic configurations is good, although there exits a tendency that energy increases are somewhat overestimated by the tight-binding methods. Serious faults have been found for the electronic structure by the tight-binding method for SiC.

1. INTRODUCTION

It is of much importance to clarify the structure and properties of complex or extended defects in semiconductors, such as planar defects, grain boundaries, and aggregates of interstitials, vacancies and impurities. By using the transferable tight-binding (TB) method and related methods [1-6], it is possible to deal with such complex defects more accurately than using previous TB methods [7] or empirical potentials [8]. However, one should be prudent about the quantitative accuracy of such semi-empirical methods applied to complex systems, because usually the accuracy of such methods has been examined only by comparison with *ab initio* results of several crystal phases or simple defects. It is desirable to compare TB calculations of complex defects directly with *ab initio* calculations. Nowadays, it is not so difficult to perform *ab initio* calculations based on the density-functional theory [9] about various defects in semiconductors if the number of atoms in the supercell is less than one or two hundred. This is owing to the development of the first-principles molecular dynamics (FPMD) method originated by Car and Parrinello [10,11]. In this paper, TB calculations of complex or extended defects such as coincidence boundaries in Si and SiC, {113} planar interstitial defects in Si, and self-interstitial clusters in Si, using the transferable TB method of Si [2,3] and the self-consistent TB (SCTB) method of SiC [12], are compared with *ab initio* calculations.

In the cases of boundaries and planar defects, it is possible to construct supercell configurations of proper size for both TB and *ab initio* methods, and a direct comparison is possible. Such comparisons should be effective for further improvement of the TB methods. In the case of self-interstitial clusters, it is desirable to contain over several hundred atoms in one unit cell, which can be dealt with only by the TB method. We apply the *ab initio* method to a configuration in a small cell so as to test some aspects of the TB results of a large cell.

2. THEORETICAL METHOD

Transferable Tight-Binding Method for Si

The transferable TB method of Si [2,3] was developed in order to overcome the weak point of the previous TB method [7] in which the transferability of structures of Si other than the fourfold coordinated ones is not necessarily guaranteed. In this method, the binding energy, which is the difference between the total energies of the system and the free atoms, is expressed as a sum of the band-structure energy E_{bs} and the remaining repulsive energy E_{rep}. The former is obtained by tight-binding band calculations with the valence atomic-orbital basis, and the latter is given as a sum of short-range repulsive interatomic potentials. As compared with the previous TB methods, this method has three difining characteristics. First, the behavior of the two-center integrals in the Hamiltonian and that of the interatomic potential are modified for large distances and are smoothly truncated by using proper attenuation functions [1]. Second, different exponent values are used for respective two-center integrals and potential, and the parameters are determined so as to reproduce the Hamiltonian including second-neighbor interactions. Third, the dependence on the local environment is incorporated into E_{rep} through the effective coordination numbers. E_{rep} is automatically reduced in the enviromnent with large coordination numbers.

As discussed in [3], the dependence on the local environment of E_{rep} is especially important for the transferability, and was incorporated for the first time. This is because the overlap interaction between basis orbitals is one of the origins of E_{rep} and has the dependence on the local environment as well as E_{bs}. This method can deal with energies and structures of various phases of Si, and is now one of the most reliable TB methods. A weak point of this method is that the bulk band-gap width of Si is overestimated as 2.2eV. In this method, the intra-atomic electrostatic interactions are included self-consistently through a form of a Hubbard-like Hamiltonian [13].

Self-Consistent Tight-Binding Method for SiC

In order to deal with defects in SiC, we developed the SCTB method [12]. This is the application of the SCTB method originally proposed by Majewski and Vogl [14]. This method can deal with the effects of self-consistent charge redistribution and electrostatic interactions within the tight-binding approximation. This is essential for defects in solids with both ionic and covalent characters. In SiC, there exists substantial charge transfer from Si toward C, which gives rise to significant ionic characters [15].

The tight-binding Hamiltonian is expressed by assuming an r^{-2} dependence on the interatomic distance r. Charge transfer effects are included self-consistently in the on-site elements of the Hamiltonian. The on-site element consists of the following terms: the orbital energy of the neutral atom, the change in the intra-atomic electrostatic potential caused by the on-site charge, the inter-atomic electrostatic potential, and the nonorthogonality correction caused by the overlap between neighboring atomic oribitals. Finally, the binding energy is expressed as a sum of the following energy terms: the promotion and intra-atomic electrostatic energy E_{pro}, the inter-atomic electrostatic energy E_{Mad}, the covalent energy E_{cov}, and the repulsive overlap energy E_{ov}. The parameters and functional forms were determined so as to reproduce the basic properties of SiC, Si and diamond. For SiC and Si, the fourfold coordinated structures are reproduced as the most stable ones. This method is one of the

few practical TB methods dealing with total energies of compounds.

Ab Initio Method

We use the plane-wave pseudopotential method based on the density-functional theory. In order to deal with large systems, we use the techniques of the FPMD method. First, the electronic minimization is performed efficiently through iterative techniques such as the conjugate-gradient method [11]. For the minimization of the total-energy functional E_{tot} with the orthonormalization constraint of wave functions $\{\psi_i\}$, we consider the functional $E = E_{tot} - \sum \lambda_{ij} (\langle \psi_i | \psi_j \rangle - \delta_{ij})$ where λ_{ij} is the Lagrange multipliers. Its gradient with respect to ψ_i can be expressed as $\delta E / \delta \psi_i^* = (H - \lambda_i) \psi_i$ where $\lambda_i = \langle \psi_i | H | \psi_i \rangle$, if the orthonormalization process is performed separately. The negative of this gradient can be regarded as a force or a steepest-descent vector acting on ψ_i. Thus the wave functions can be iteratively updated according to this force or vector and can reach the true eigen functions minimizing E_{tot}. Thus one can avoid time-consuming computations of the direct diagonalization of the Hamiltonian.

Second, we use optimized pseudopotentials by Troullier and Martins [16] so as to reduce the number of plane-wave basis functions N_{pw}. The TM pseudopotentials were constructed by using the configuration $3s^2 3p^{0.5} 3d^{0.5}$ for Si and $2s^2 2p^2$ for C. The cutoff radius is 2.0a.u. for the s, p and d orbitals of Si, and 1.44a.u. for the s and p orbitals of C. We use the separable form [17] with the p pseudopotential as the local component for both elements. The lattice constants, cohesive energies and bulk moduli of bulk Si, SiC and diamond are reproduced accurately within the local density approximation [18] by using the plane-wave cutoff energy E_{cut} of 35Ry for Si and of 60Ry for SiC and diamond. The equilibrium lattice constants of Si and SiC are 98.9% of the experimental values of 0.5429nm and 0.4360nm. That of diamond is 99.3% of the experimental value of 0.3567nm. All calculations of defects are performed by using these equiliburium lattice constants. We also use the Hamman-Schlüter-Chiang (HSC) pseudopotential of Si [19], because this potential seems to provide more accurate results for small E_{cut} of 30Ry. In this case, the equilibrium lattice constant of Si is the same value. It should be noted that the total energies are well converged for the above values of E_{cut} for respective pseudopotentials within the accuracy of 0.02eV/atom as compared with the values for E_{cut} of 60Ry for Si and 100Ry for SiC and diamond.

3. GRAIN BOUNDARIES IN Si

Grain boundaries in Si significantly affect the properties of polycrystalline Si used for solar cells and various devices such as thin-film transistors. Thus it is of much importance to understand the atomic and electronic structures of grain boundaries. Various tilt and twist boundaries in Si were dealt with by using the transferable TB method [20,21]. It has been pointed out that the $\langle 011 \rangle$ tilt coincidence boundaries are especially stable as compared with twist or general boundaries. This is because stable configurations can be easily constructed for such boundaries by arranging structural units consisting of atomic rings without any dangling bonds or large bond distortions. This point is in good agreement with experimental observations using high-resolution transmission electron microscopy (HRTEM) [22]. TB calculations were also used to clarify the effects of structural disorder at grain boundaries such as bond distortions, odd-membered rings and coordination defects on the electronic structure, which provided useful information about the origin of the gap states and band tails at grain boundaries in Si [21].

We apply the *ab initio* method to the $\{122\}\Sigma = 9$ boundary in Si. This is a typical tilt boundary. We deal with the glide-plane model, where the interface is constructed by zigzag arrangement of five- and seven-membered rings without any dangling bonds. We use a 64-atom cell. The ridid-body translation optimized by the TB method [20] is used for the initial configuration. We use the TM pseudopotential with E_{cut} of 35Ry. Two special **k** points per irreducible eighth of the Brillouine zone are used. After relaxation, the energy increase against the bulk crystal is calculated by using four special **k** points for the relaxed configuration in order to obtain a more accurate value. This is because the energy increase of this boundary is rather small, which is about 1eV per cell. To execute heavy calculations with four **k** points, it is desirable to reduce E_{cut} without the loss of accuracy. Thus we have used the HSC pseudopotential with E_{cut} of 30Ry as mentioned in Sec. 2. We have found that the relaxed configuration has negligible atomic forces for this new condition.

(a)   (b)
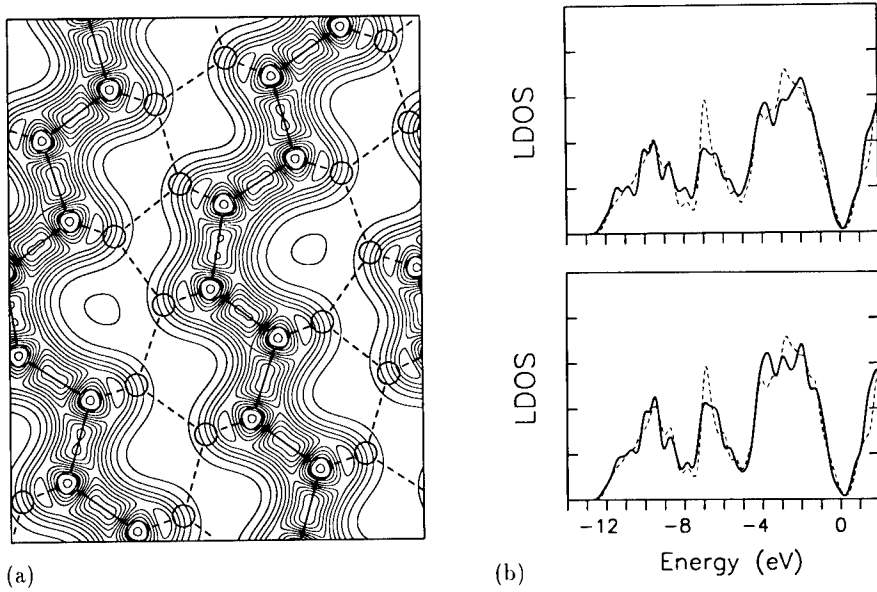
Figure 1. (a) Relaxed configuration and valence electron density of the $\{122\}\Sigma = 9$ boundary in Si. Contours of the electron density are plotted from $0.003a.u.^{-3}$ to $0.091a.u.^{-3}$ in spacing of $0.008a.u.^{-3}$. Circles indicate the atomic position projected along the $\langle 011 \rangle$ axis. (b) Local densities of states (LDOS's) of this boundary. The LDOS's of the upper and lower panels are those of the interface and bulk regions, respectively. Each region corresponds to one half of the 64-atom cell. The dashed line is the DOS of diamond Si. Six **k** points in the irreducible part of the Brillouine zone and the Gaussian broadening of $\sigma = 0.2eV$ are used.

Table 1 lists the results of energy and bond distortions. Figure 1(a) shows the relaxed configuration and charge distribution. It is clear that the present configuration is stable with small bond distortions and energy increase. All the interfacial bonds are well reconstructed, although stretched bonds have slightly reduced bond charges.

Table 1. Calculated results of the $\{122\}\Sigma=9$ boundary in Si. SW indicates results by the SW potential [8]. E_{gb} is the boundary energy, which is the energy increase per unit area of one interface against the bulk crystal. Δr and $\Delta\theta$ are the bond-length and bond-angle distortions.

	$E_{gb}(J/m^2)$	$\Delta r(\%)$	$\Delta\theta(degree)$
ab initio (64-atom cell)	0.18	–1.8 to +1.4	–16.3 to +22.2
TB (144-atom cell)	0.32	–1.9 to +1.5	–16.0 to +19.9
SW (144-atom cell)	0.45	–1.5 to +1.9	–16.3 to +21.1

The agreement between the TB results and the *ab initio* results is rather good, especially about the atomic configuration. The calculated stress [23] in the direction normal to the interface is rather small (-0.5GPa), which justifies the initial translation normal to the interface from the TB results. There is a substantial difference in the energy value. The TB method overestimates the energy increase as does the SW potential. Of course, there exist the problems of cell size and of energy convergence associated with the number of **k** points, N_{pw} and E_{cut} [24]. However, the same tendency is observed in other cases.

We have examined the electronic states of several **k** points. There exist no electronic states inside the band gap in accordance with reconstruction, although there exist states with relatively large probabilities at the interface region near the band edges. This feature is similar to the TB results [21]. As shown in Figure 1(b), it does not seem that the present interface has a large effect on the electronic structure. However, there exist small changes in the shape of the valence band DOS. The *s-p* mixing peak is decreased and the states are increased at the valleys among the three peaks. The *p-p* bonding peak tends to be flattened. These features are also similar to the TB results [21] and seem to be caused by the presence of odd-membered rings.

4. {113} PLANAR DEFECTS IN Si

The {113} planar or rodlike defects are induced by electron irradiation, ion implantation, or thermal anealing only in Si or Ge. These defects are considered as aggregations of supersaturated self-interstitials. However, it is very strange that interstitial atoms are aggregated not on simple planes but on {113} planes. This difficult problem was solved for the first time by modelling through HRTEM observation coupled with theoretical calculations [25-28]. The model is described by an arrangement of two kinds of structural units, I and O units, on the {113} plane [25]. These units consist of atomic rings without any dangling bonds on the $\langle 110 \rangle$ projection, similar to those in the $\langle 110 \rangle$ tilt boundaries shown in Sec. 3. An I unit contains a self-interstitial atom chain along the $\langle 110 \rangle$ direction. The core of this unit corresponds to a tiny rod of hexagonal Si surrounded by five- and seven-membered rings. An O unit is formed from an I unit by removing the self-interstitial chain, which leaves an eight-membered ring. In the observed sequence of units along the $\langle 332 \rangle$ direction in Si [25], the ratio of I units is about 62% and there exists no long-range periodicity. However, there exists some short-range order. O units are arranged separately at intervals of one or two I units, occasionally three I units.

On the theoretical side, calculations using the SW potential showed that the model has relatively small bond distortions and energies [26]. The relaxed configuration is in good

agreement with the HRTEM image [27]. The calculations also indicated that I units are more stable when neighboring side by side on the {113} plane rather than isolated I units, and that the insertion of O units can stabilize the sequence of I units. From these points, the mechanism of generation and the origin of the observed arrangement of the units were deduced [26,27]. These results were supported by the TB calculations of the $|IO|$ and $|IIO|$ models [28]. In these models, the structural units sandwiched by "|" are repeated periodically along the $\langle 332 \rangle$ direction. It is possible to investigate essential features of the real structure with only short-range order by such periodic models.

In this paper, we deal with the $|IO|$ model by the *ab initio* method. We use the same 90-atom cell used in the previous calculations [26,28]. In the unit cell, one sequence of the units is contained between the bulk regions with the observed rigid-body translation [25]. We use the TM pseudopotential and E_{cut} of 35Ry. Two special **k** points are used. After the ralaxation, the energy increase is calculated more accurately by using four **k** points with the HSC pseudopotential and E_{cut} of 30Ry because of the same reason mentioned in Sec. 3.

Table 2 lists the results. The agreement between the TB and *ab initio* results is rather good. It is clear that the TB results are better than the SW results, which is also seen in the results of Sec. 2. The tendency to overestimate energy increases in the TB and SW calculations is also observed.

Table 2. Calculated results of the $|IO|$ model of the {113} planar intersitial defects in Si. E_{is} is the energy increase per interstitial atom. This model contain two interstitial atoms per unit cell. Δr and $\Delta \theta$ are the bond-length and bond-angle distortions. All calculations use the same 90-atom cell.

	$E_{is}(eV/atom)$	$\Delta r(\%)$	$\Delta\theta(degree)$
ab initio	0.49	−1.7 to +3.0	−17.6 to +23.7
TB	0.92	−1.8 to +2.8	−17.9 to +24.0
SW	1.13	−1.4 to +2.6	−18.8 to +22.4

Figure 2 shows the relaxed configuration and valence electron density. It is clear that all the bonds are well reconstructed, although distorted bonds have reduced bond charges. It can be said that the *ab initio* calculation has proved the stability of the $|IO|$ model and the stability of the structural units. It is important to deal with the $|IIO|$ model so as to examine the relation between the two models, which is now in progress. Previous calculations indicated the relative stability of the latter model. This is reasonable because the latter model more closely resembles the observed sequence of the units.

The electronic structure of the $|IO|$ model has been examined. This configuration has electronic structure similar to that of the $\langle 011 \rangle$ tilt boundary in Sec. 2. In accordance with reconstruction, there exist no states inside the band gap, although we found states with large probabilities associated with the structural units near the band edges. These features are similar to the TB results [28].

Recently, high-resolution transmission electron energy-loss spectroscopy (HR-TEELS) has been applied to {113} defects in Si [29]. The obtained low-loss and core-loss spectra with energy resolution of 0.1eV have indicated that these kinds of defects at least contain no deep states in the band gap. This point supports our prediction. A remarkable change by the defects has been observed in the Si-$L_{2,3}$ ELNES (electron energy-loss near edge structure),

which corresponds to s and d-projected densities of states of conduction band. The first peak of the conduction band, mainly consisting of s-like states, is sharpened by the decrease of the higher portion of this peak. We have calculated the LDOS's of the $|IO|$ model in order to explain the observed change in the ELNES. The LDOS's have similar features to those of the $\Sigma = 9$ boundary in Figure 1(b). In the valence band, remarkable changes are the decrease of the s-p mixing peak, the increase of states at the valleys among the three peaks, and the flattening of the p-p bonding peak. In the conduction band, two sharp peaks are also flattened. These seem to be caused by odd-membered rings and bond distortions [30]. The observed change in the ELNES has been explained by the decrease of the s-state densities at the higher portion of the first peak caused by odd-membered rings [29].
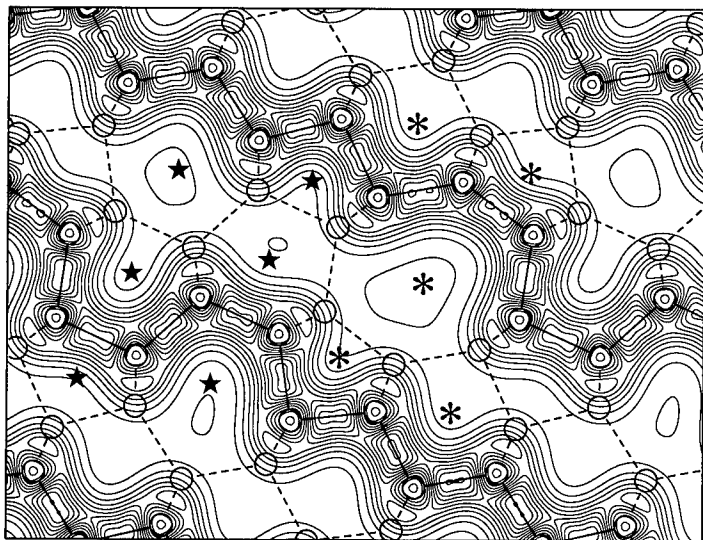
Figure 2. Relaxed configuration and valence electron density of the $|IO|$ model of the $\{113\}$ planar interstitial defects in Si. Contours of the electron density are plotted from $0.003a.u.^{-3}$ to $0.091a.u.^{-3}$ in spacing of $0.008a.u.^{-3}$. Atomic rings constituting I and O units are indicated by stars and asterisks, respectively.

5. SELF-INTERSTITIAL CLUSTERING IN Si

Self-interstitial atoms play an important role in various dynamic phenomena in Si such as self-diffusion and impurity diffusion. However, the existence of self-interstitial atoms is confirmed experimentally only when they form very extended defects such as $\{113\}$ planar defects. Thus it is of great importance to investigate structure and properties of primary interstitial defects and interstitial clusters. In contrast to lattice vacancies and their complexes, it is very strange that self-interstitials and their clusters seem to generate no active electronic states in the band gap nor optical centers. Electron paramagnetic resonance has yielded no convincing structural information about defects of self-interstitials. Therefore, the structures of self-interstitials and their clusters have always been under debate, except

for very extended defects. Interstitial clusters of small or intermediate size are also of much interest so as to understand the mechanism of formation of the {113} planar defects.

No clear experimental signals of self-interstitials and their clusters imply the possibility that self-interstitials usually exist as agglomarates without any coordination defects. In the {113} interstitial defects, lattice reconstruction occurs in the interior of the Si crystal involving self-interstitials so as to remove dangling bonds as shown in Sec. 4. Thus we have proposed a novel model of self-interstitial clustering in Si where all the atoms are fourfold coordinated [31]. Figure 3 shows the model of the four interstitial cluster I_4. This cluster contains four self-interstitials. The configuration is characterized by succesive five- and seven-membered rings without any coordination defects.

The stability of this cluster has been examined by the SW potential and the TB method. We have used the 724-atom cell containing one I_4 cluster. The present structure seems to induce large bond distortions and strains. Thus it is desirable to use cells of large enough size. The number of atoms in the cell is inevitably large as compared with cells for coincidence boundaries or planar defects with some periodic features. The edge lengths of the 724-atom cell have been optimized by iterating relaxations by the SW potential so as to reduce long-range stresses. In the final cell, the cluster repeats with intervals of 2.30nm, 2.30nm and 2.73nm along the the $\langle 011 \rangle$, $\langle 0\bar{1}1 \rangle$ and $\langle 100 \rangle$ directions, respectively. Then relaxation has been performed by the TB method. Results are shown in Figure 3 and Table 3.
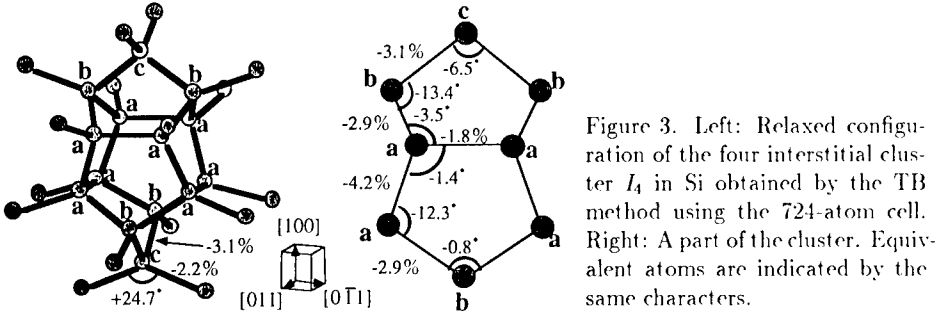
Figure 3. Left: Relaxed configuration of the four interstitial cluster I_4 in Si obtained by the TB method using the 724-atom cell. Right: A part of the cluster. Equivalent atoms are indicated by the same characters.

Table 3. Calculated results of the four interstitial cluster I_4 in Si. E_{is} is the energy increase per interstitial atom. Δr and $\Delta \theta$ are the bond-length and bond-angle distortions. The *ab initio* method is used only for the energy calculation of the relaxed configuration by the TB method.

	$E_{is}(eV/atom)$	$\Delta r(\%)$	$\Delta \theta(degree)$
TB (724-atom cell)	2.46	−4.2 to +1.3	−22.3 to +24.7
SW (724-atom cell)	2.10	−5.0 to +1.5	−22.2 to +22.7
ab initio (196-atom cell)	1.53		
TB (196-atom cell)	2.34	−3.9 to +2.1	−22.4 to +23.5

We apply the *ab initio* method to this model. We use a 196-atom cell. It is at least possible to obtain information about the stability of the present model, although this small cell may affect some aspects of the results. We use the HSC pseudopotential with E_{cut} of

294

20Ry. Only the Γ point is used. The edge lengths of the cell are optimized by the SW potential. The cluster repeats with intervals of 1.54nm, 1.54nm and 1.67nm along three directions, respectively. Relaxation has been performed by the TB method. Then the *ab initio* method has been applied to this configuration. The energy increase has been obtained as compared with the *ab initio* results of the 192-atom cell without any self-interstitials by similar conditions. All the results are listed in Table 3. Full relaxation by the *ab initio* method is now in progress.

The TB results of the 724-atom cell indicate the stability of this model. The calculated energy values are lower than those of an isolated self-interstitial atom in the range from about 3 to 7 eV by the *ab initio* or TB methods [32]. And the ranges of the bond-length and bond-angle distortions are not so different from the results of $\langle 011 \rangle$ tilt boundaries and $\{113\}$ defects. The TB results of the 196-atom cell reveal similar features, although the energy is lowered by the interaction among defects. The energy value by the *ab initio* method is smaller than the value of the TB method. This value is also much smaller than the *ab initio* values of an isolated interstitial [32]. Of course, relaxation by the *ab initio* method has not been performed. However, the atomic forces by the *ab intio* method for the present configuration are not large. Thus it can be said that the *ab initio* calculation supports the stability of the I_4 model.

About the electronic struture of the I_4 model, the TB calculation of the 724-atom cell reveals that this structure contains no deep states inside the band gap. However, defect-localized states have been found. These are occupied shallow states at the top of the valence band, and states at the bottom of the valence band. These states tend to possess a higher amplitude at the atoms with shortened bonds. This defect contains many shortened bonds because of insertion of interstitials. The origin of such states can be explained by the widening of the valence band width by bond shortening. Such electronic states will be examined by the *ab initio* method in the near future.

6. GRAIN BOUNDARIES IN SiC

SiC is a very important material for high-performance ceramics and high-temperature devices. It is of great importance to understand grain boundaries in SiC, because various properties of SiC ceramics are dominated by grain boundaries. As compared with boundaries in Si, there exist further complexities in grain boundaries in SiC, such as polarity, interface stoichiometry and wrong bonds between like atoms. We have dealt with several coincidence tilt boundaries in SiC by using the SCTB method [33].

In this paper, we apply the *ab initio* method to the $\{122\}\Sigma = 9$ boundary in SiC. HRTEM observations of this boundary [34] indicate that this boundary has a configuration with a bond network similar to the same boundary in Si in Sec. 3. However, it is possible to construct three kinds of models by inverting the polarity of grains. These are two kinds of polar (nonstoichiometric) interfaces and a non-polar (stoichiometric) interface [33]. The interface stoichiometry dominates the kinds and numbers of wrong bonds. The models of the $\Sigma = 9$ boundary have four interfacial bonds per period as shown in Sec. 3. In the N-type polar interface, two bonds in the four interfacial bonds are C-C bonds, and the others are usual Si-C bonds. In the P-type polar interface, two bonds in the four interfacial bonds are Si-Si bonds. In the non-polar interface, two bonds are both C-C and Si-Si bonds.

We use a 64-atom cell containing two interfaces in each cell. The supercell of the polar interfaces must contain two inequivalent polar interfaces by the crystallographic constraint.

Initial configurations are constructed by introducing the rigid-body translations between bulk regions otpimized by the SCTB method [34]. We use the TM pseudopotentials and E_{cut} of 60Ry. Two special **k** points are used. Results are shown in Table 4 and Figure 4.

Table 4. Calculated results of the polar and non-polar interfaces of the $\{122\}\Sigma=9$ boundary in SiC. ΔE_{tot} is the energy increase per cell against bulk SiC. E_{gb} is the boundary energy, which is obtained only for the non-polar interface. Δr and $\Delta\theta$ are the bond-length and bond-angle distortions. Ranges of Δr include only distortions of Si-C bonds. Values of Δr for C-C and Si-Si bonds mean the deviations against the bond lengths in bulk diamond and bulk Si, respectively.

	$\Delta E_{tot}(eV/cell)$	$E_{gb}(J/m^2)$	$\Delta r(\%)$	$\Delta\theta(degree)$
Non-Polar				
ab initio	4.44	1.27	-2.9 to +2.9	-22.4 to +27.9
(64-atom cell)			C-C:+3.6, Si-Si: 4.4	
SCTB	5.07	1.43	-2.5 to +2.2	-23.1 to +24.1
(144-atom cell)			C-C:+3.8, Si-Si: 4.3	
Polar				
ab initio	4.31			
N-type			-2.7 to +2.0, C-C:+4.5	-20.1 to +22.5
P-type			-2.7 to +2.5, Si-Si:-1.1	-13.0 to +24.0
SCTB	4.52			
N-type			-2.5 to +2.1, C-C:+2.5	-21.8 to +19.7
P-type			-2.1 to +1.4, Si-Si: 1.9	-13.1 to +21.1

It can be said that all the interfacial bonds are well reconstructed in each interface. The bond lengths and bond charges of the C-C and Si-Si bonds in each interface have features similar to those in bulk diamond and Si. The C-C bonds in the non-polar and N-type polar interfaces has a two-peak structure in the charge density, typical of bulk diamond. It seems that the electrostatic repulsion between C atoms with partial negative charges lengthens the C-C bond as compared with bulk diamond. The agreement between the *ab initio* and TB results is rather good on the atomic configurations, except for the C-C bonds at the N-type polar interface. The agreement on the total energy is also good, although the difference between the non-polar and polar interfaces is larger in the TB results. The relative stability of the polar interfaces can be explained by the relatively small bond distortions because of only one kind of wrong bonds in each polar interface.

As shown in Figure 4(b), the present interfaces in SiC have a clear band gap in accordance with reconstruction of interfacial bonds. Electronic states greatly localized at the C-C bonds are observed at the top and bottom of the valence band, which are indicated by stars in Figure 4(b). The localized states at the valence-band top exist at about 0.3eV and 0.5eV above the bulk valence-band top for the non-polar and N-type polar interfaces, respectively. States weakly localized at the Si-Si bonds are also observed near the bulk valence-band top for the non-polar and P-type polar interfaces, which are indicated by asterisks in Figure 4(b).

The origin of the C-C bond states can be explained by the analogy to a diamond cluster embedded in bulk SiC. The width of the valence band of diamond is much larger than that of SiC. Thus band-edge states of diamond should stick out from the SiC valence band, and

become localized states. The present C-C bond states at the band bottom consist of s-like orbitals of C atoms, and the C-C bond states above the bulk band top consist of p-like orbitals of C atoms, similarly to the bulk band-edge states of diamond.
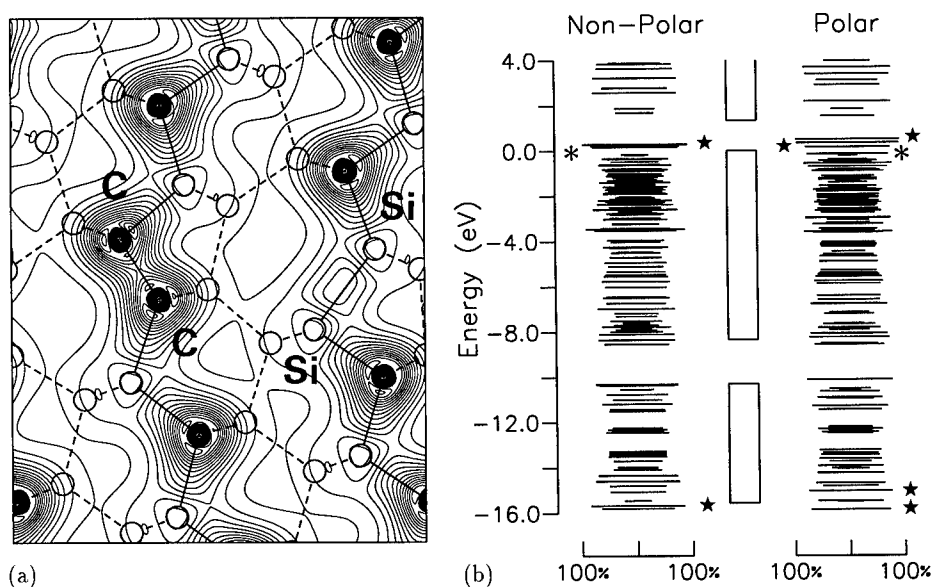
Figure 4. (a) Relaxed configuration of the non-polar interface of the $\{122\}\Sigma = 9$ boundary in SiC. Contours of the valence electron density are plotted from 0.015 to 0.295 in spacing of 0.020 for SiC. (b) Eigen energy levels at the Γ point for the polar and non-polar interfaces. Length of each horizontal line represents the probability of each state in the interface regions, which correspond to one half of the unit cell. Boxes represent the bulk band projection.

In the results by the SCTB method [34], the C-C bonds generate localized states only at the bottom of the valence band, and the Si-Si bonds generate localized states at the top of the valence band and the bottom of the conduction band. This is a serious fault of the SCTB method. This may be caused by the usage of the universal parameters and simple r^{-2} dependence for the two-center integrals [14].

7. CONCLUSION

Essential results of our TB calculations of various extended defects in Si and SiC have been supported by *ab initio* calculations. Especially, stable configurations by the TB methods are rather correct, although the energy values are often overestimated and the C-C bond states in SiC are not correct. The present results also indicate the importance of an effective combination of the TB method and the *ab initio* method in dealing with very complex systems, as is now being executed for the self-interstitial clusters in Si.

ACKNOWLEDGMENTS

REFERENCES

[1] L. Goodwin, A.J. Skinner and D.G. Pettifor, Europhys. Lett. **9**, 701 (1989).
[2] S. Sawada, Vacuum **41**, 612 (1990).
[3] M. Kohyama, J. Phys. Condens. Matter **3**, 2193 (1991).
[4] I. Kwon et al., Phys. Rev. B **49**, 7242 (1994).
[5] C.Z. Wang et al., MRS Symp. Proc. **408**, 37 (1996).
[6] M. Menon and K.R. Subbaswamy, Phys. Rev. B **55**, 9231 (1997).
[7] D.J. Chadi, Phys. Rev. Lett. **41**, 1062 (1978); Phys. Rev. B **29**, 785 (1984).
[8] F.H. Stillinger and T.A. Weber, Phys. Rev. B **31**, 5262 (1985).
[9] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L.J. Sham, *ibid.* **140**, A1133 (1965).
[10] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[11] M.C. Payne et al., Rev. Mod. Phys. **64**, 1045 (1992).
[12] M. Kohyama et al., J. Phys. Condens. Matter **2**, 7791 (1990).
[13] D. Tomanek and M.A. Schlüter, Phys. Rev. Lett. **56**, 1055 (1986).
[14] J.A. Majewski and P. Vogl, Phys. Rev. B **35**, 9666 (1987).
[15] K.J. Chang and M.L. Cohen, Phys. Rev. B **35**, 8196 (1987).
[16] N. Troullier and J.L. Martins, Phys. Rev. B**43**, 1993 (1991).
[17] L. Kleinman and D.M. Bylander, Phys. Rev. Lett. **48**, 1425 (1982).
[18] J.P. Perdew and A. Zunger, Phys. Rev. B**23**, 5048 (1981).
[19] D.R. Hamman, M. Schlüter and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979); M.T. Yin and M.L. Cohen, Phys. Rev. B **26**, 5668 (1982).
[20] M. Kohyama and R. Yamamoto, Phys. Rev. B **49**, 17102 (1994).
[21] M. Kohyama and R. Yamamoto, Phys. Rev. B **50**, 8502 (1994).
[22] Y. Ishida and H. Ichinose, in *Polycrystalline Semiconductors*, edited by H.J. Möller, H.P. Strunk and J.H. Werner (Springer, Berlin, 1989), p.42.
[23] O.H. Nielsen and R.M. Martin, Phys. Rev. B **32**, 3780 (1985); **32**, 3792 (1985).
[24] G.-M. Rignanese et al., Phys. Rev. B **52**, 8160 (1995).
[25] S. Takeda, Jpn. J. Appl. Phys. **30**, L639 (1991).
[26] M. Kohyama and S. Takeda, Phys. Rev. B **46**, 12305 (1992).
[27] S. Takeda, M. Kohyama and K. Ibe, Phil. Mag. A **70**, 287 (1994).
[28] M. Kohyama and S. Takeda, Phys. Rev. B **51**, 13111 (1995).
[29] H. Kohno, T. Mabuchi, S. Takeda, M. Kohyama, M. Terauchi and M. Tanaka, submitted.
[30] J. Singh, Phys. Rev. B **23**, 4156 (1981).
[31] N. Arai, S. Takeda and M. Kohyama, Phys. Rev. Lett. **78**, 4265 (1997).
[32] P.J. Kelly and R. Car, Phys. Rev. B **45**, 6543 (1992); C.Z. Wang et al., Phys. Rev. Lett. **66**, 189 (1991).
[33] M. Kohyama et al., J. Phys. Condens. Matter **2**, 7809 (1990); **3** 7555 (1991); MRS Symp. Proc. **339**, 9 (1994).
[34] K. Hiraga, Sci. Rep. Res. Inst. Tohoku Univ. A**32**, 1 (1981).

EMPIRICAL TIGHT-BINDING APPLIED TO SILICON NANOCLUSTERS

G. ALLAN, C. DELERUE, M. LANNOO
Institut d'Electronique et de Microélectronique du Nord, Département Institut Supérieur d'Electronique du Nord, BP 69, 59652 Villeneuve d'Ascq Cedex, France, gal@isen.fr

ABSTRACT

The calculation of the electronic structure of silicon nanostructures is used to discuss the accuracy of results obtained by the tight-binding method. We first show that the level of refinement of the tight-binding approximation must be adapted to the calculated property. For example, an accurate description of both the valence and conduction bands which can be achieved with a 3rd-nearest neighbor approximation is necessary to calculate the variation of the gap energy with the silicon crystallite size. The sp3s* model which gives a bad description of the conduction band underestimates the confinement energy but can give good results when it is used to determine the variation of the crystallite band gap with pressure. To study Si-III (BC-8) nanocrystallites, we show that a good description of the bulk band structure can be obtained with non-orthogonal tight-binding but due to the large number of nearest neighbors one must take analytical variations of the parameter with interatomic distances. The parameters involved in these expressions can be easily fitted to the bulk band structures using the k-point symmetry without requiring the use of group theory. Finally we discuss the effect of increasing the size of the minimal-basis set and we show that it would be possible to get the values of the tight-binding parameters from a first-principles localized states band structure calculation avoiding the fit to the energy dispersion curves.

INTRODUCTION

The discovery of porous silicon luminescence [1] has induced many theoretical studies of silicon nanocrystallites properties. The blue shift of a cluster gap which is due to quantum confinement can be simply explained in the effective mass approximation (EMA) but it has been rapidly shown that it is overestimated by this method for the experimental cluster sizes (~1-10 nm). Empirical methods (using pseudopotentials or tight-binding (TB)) seem to be well suited for such sizes filling the gap between EMA and first-principles method.

In the next section, we show that an accurate description of the valence and conduction bands is necessary to get accurate variations of the blue shift. This is the case of EMA or of TB approximations limited to first or second nearest interactions which respectively overestimate or underestimate the cluster band gap. We also show that a TB model can give good results for the blue shift but also can give very bad ones for other properties like the deformation potential.

The third section is devoted to the fit of empirical TB parameters to a known band structure calculated by a first principles method like the pseudopotential local density approximation for example. Applied to the Si-III (BC-8) phase, we use simple analytic expression for the TB parameters. Such expressions reduce the number of parameters one has to fit. We also show how symmetry can be easily used even with such a large unit cell.

BLUE SHIFT OF SILICON CLUSTER GAPS

This is the first property which has been computed as it can be compared with the experimental luminescence energy. The EMA applied to a spherical infinite well gives a d^{-2} variation of the gap with the diameter d of the cluster. It is obvious that such a variation cannot be valid for very small clusters (for very small clusters, the gap diverges). For large clusters TB can give the same
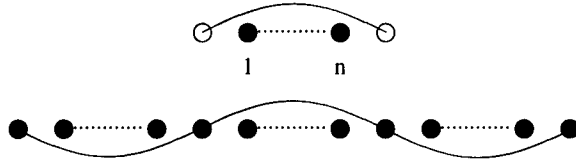
Fig. 1 Eigenstate of a linear cluster compared to the eigenfunction of the linear chain/ for the wave vector equal to $\frac{\pi}{(n+1)a}$

kind of results provided the effective masses in this approximation are in good agreement with the experimental values. This is not the case of the sp3s* Vogl's model [2]. The width of the conduction band is underestimated and the cluster gaps calculated with this model are too small.

A simple explanation can be obtained from a TB model applied to a linear chain of atoms (Fig. 1). Periodic clusters of n atoms can be obtained by removing one atom each n+1 sites. This is sufficient with TB parameters limited to nearest neighbors to suppress all hamiltonian matrix elements between the clusters. Creating such vacancies can be simulated in TB by a local infinite potential on the atoms which have to be removed [3]. The effect of such a potential is to cancel the component of the linear chain eigenstates on these states. So the extended states of the clusters are simply eigenstates of the linear chain with a component equal to zero on the vacancies. If $-\beta$ is the TB nearest neighbor interaction, the energy E_n of the cluster is equal to $-2\beta\cos(ka)$ where a is the lattice parameter and $k = \frac{l\pi}{(n+1)a}$ $l = 1, 2, ...n$ so that the linear chain eigenstate $\psi \sim \sin(kx)$ vanishes on each vacancy site $x = (n+1)a$. Such truncated crystal model has been used for one-orbital per atom representations [4], single-band models [5-8] and empirical pseudopotentials [9-10]. However this only gives extended clusters or slabs states, localized states as dangling bonds for example can be obtained only if one considers complex wave vectors.

Such a method could be generalized to a 3 dimensional cluster provided one uses a TB model without three-centers parameters. In such a case one must also suppress the contribution due to potential localized on the removed atoms. The cluster wave function can be expressed as a linear combination of bulk states with wave vector k:

$$\Psi_n = \sum_k a_{k,n}\Psi_k \tag{1}$$

corresponding to an energy belonging to the bulk band energy range:

$$E_n = \sum_k a_{k,n}^2 E_k \tag{2}$$

A non-orthogonal TB model [11] with a minimal sp3 basis gives a good description of both the valence and conduction bands. We have already noticed that this is not the case of orthogonal TB models like the sp3s* one which gives very bad effective masses. The agreement for the lowest conduction bands with first principles calculations can be improved with next-nearest TB [12] and for higher energy bands one has to consider third-nearest and three-centers TB interactions [13] or take a minimal basis with s, p and d atomic functions. We have applied these different models to spherical silicon nanocrystallites. The dangling bonds on the surface atoms are saturated by hydrogen atoms. The blue shift is given on Fig. 2. One can see the quite large difference between the different models for small clusters and which is mainly due to the width of the bulk conduction band. Close results are obtained (Fig. 2 curves a and b) for the two TB models which give a bulk conduction band in very good agreement with the first-principles one.

Our best values obtained with non-orthogonal TB [14] are also in very good agreement with the empirical pseudopotentials ones [15] and the first-principles ones [16] (Fig. 3). Let us notice that in the latter method, a constant correction to the LDA gap equal to the bulk one has been used
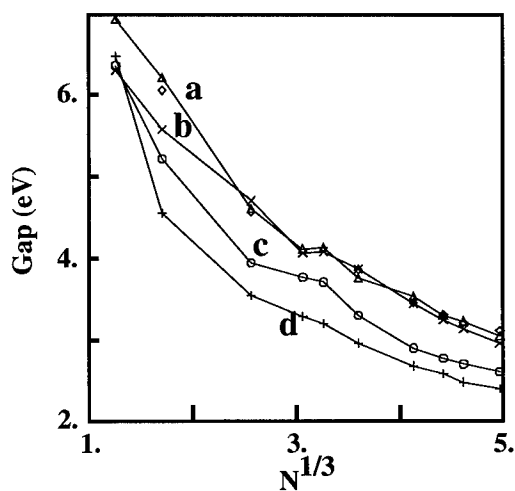
Fig. 2 Comparison between different ETB models for Si clusters passivated with hydrogen. a) Non-orthogonal tight binding parameters [11]. b) Third nearest neighbors parameters [13]. c) Second nearest neighbors parameters [12]. d) First nearest neighbors sp^3s^* parameters [2].
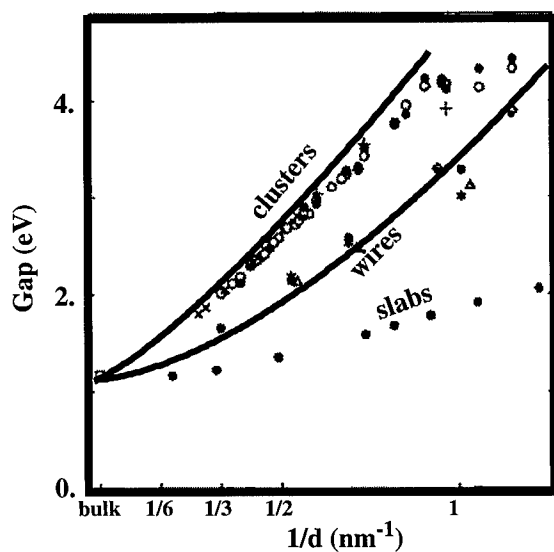
Fig. 3 Energy gap vs confinement parameter 1/d for hydrogen terminated Si clusters, wires and slabs ($d = aN/8$ for slabs, $d = a(0.5N/\pi)^{1/2}$ for wires, $d = a(0.75N\pi)^{1/3}$ where $a = 0.54nm$ and N is the number of Si atoms in the unit cell). [15]

301
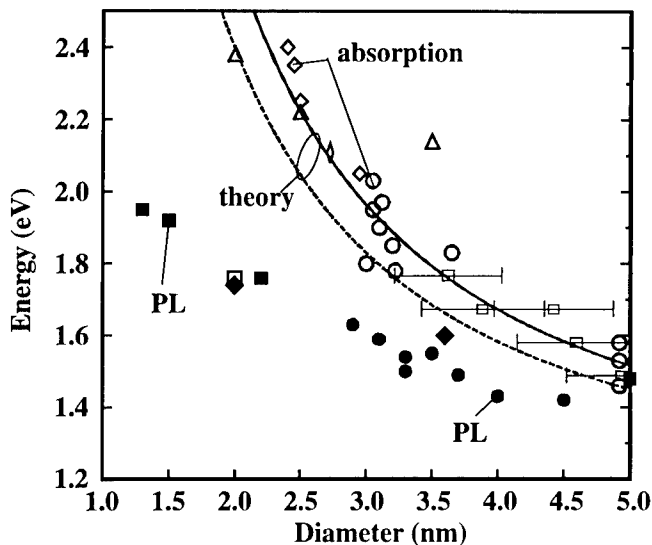
Fig. 4 Compilation [17] of optical bandgaps of silicon crystallites and porous silicon samples obtained from optical absorption (empty symbols) and luminescence (full symbols) and with luminescence of hydrogenated clusters ((□) [18]. Straight line: one-electron bandgap calculated for spherical silicon crystallites [14]. Dashed line: the same but including the excitonic binding energy

as its variation with cluster size is unknown. There is a good agreement of the computed values with the measurements made for hydrogenated clusters and with the absorption edge (Fig. 4). Nevertheless the calculated values are much larger than the luminescence energy measured for porous silicon. Such disagreement is also found for oxidized clusters and seems to be due to a surface effect [19] like self-trapped exciton [20] for example.

Analytic expressions of the non-orthogonal TB parameters as a function of the interatomic distance are given in [11] and one can try to use these values to calculate the variation of the bulk and cluster gaps with pressure. However the overlaps in this model are large This gives rise to a too large repulsion between the atoms and then to a bulk lattice parameter larger than the experimental one [21]. This also gives rise to very small eigenvalues of the overlap matrix as shown on Fig. 5. A small variation of the interatomic distance can make some of these eigenvalues to vanish and even become negative for a few k points in the Brillouin zone which leads to energy divergences. Applied to clusters, the contribution of these k points to the variation of the energy given by equation (2) with pressure is not negligible and the result does not agree for small clusters with first principles results [21]. Such discrepancy is due to the fit of the overlaps to the band structure. In the case of a minimal sp3 basis they are forced to reproduce an effect which is due to higher energy states like the d ones.

Such divergences do not occur with the orthogonal sp3s* TB model. If one uses the bulk deformation potentials measured for some high symmetry k points, one can fit the variation of the TB parameters with distance. In this model the negative indirect gap deformation potential is due to the increase repulsion between the s* state (which simulates the effect of higher energy d states) with the first conduction band with increasing pressure. We have assumed that the interatomic

302
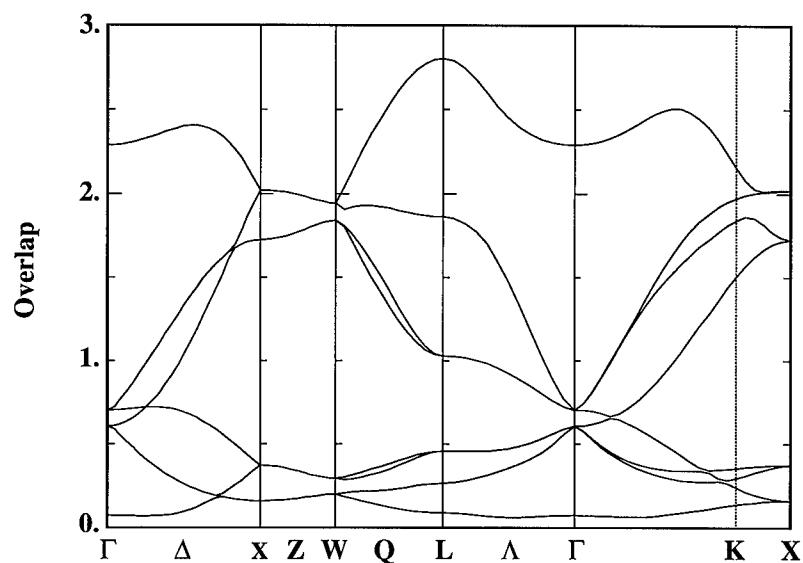
Fig. 5 Overlap dispersion curves along symmetry lines of the Brillouin zone.
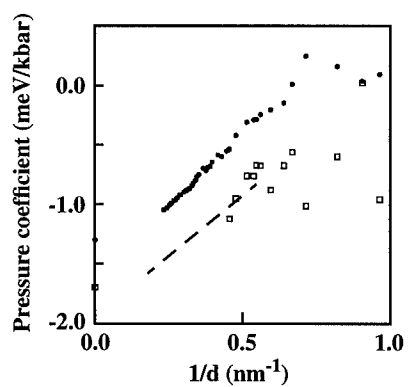
Fig. 6 Band gap pressure coefficient of spherical nanocrystallites [22] as a function of their diameter: LDA calculation [23] for small clusters (□), sp^3s^* tight-binding model (•)[2]. The values for 1/D=0. are the bulk values. The dotted line is a guide four the eyes between the LDA bulk and small clusters values.

nearest-neighbor Slater-Koster tight-binding parameters used in the Vogl model vary as d^{-n_α} where α stands for ssσ, spσ, ppσ, ppπ and s*pσ. The n_α's have been fitted to the four bulk pressure coefficients listed above. We get $n_{ss\sigma} = 3.25$, $n_{sp\sigma} = 2.5$, $n_{pp\sigma} = 1.5$, $n_{pp\pi} = n_{pp\sigma}$ and $n_{s^*p\sigma} = 3$ [22]. The comparison between TB and LDA results is given on Fig.6 as a function of cluster size. The results are slightly scattered for small clusters and one can see from the tight-binding calculation that it smoothly tends towards the bulk value when the diameter increases. For small clusters, the difference is comparable to the one between tight-binding and LDA bulk values. The absolute difference between results obtained by the two methods is quite small as bulk deformation potentials vary in the ~0-10 meV/kbar range. Such agreement suggests that the expansions given by equ. (1) of the cluster HOMO and LUMO states obtained by the two methods are quite close. The net difference between the behavior of the gap value and its variation with pressure calculated by TB and LDA methods can be simply explained with equation (2). This is due to the good agreement between the bulk deformation potentials calculated by the two methods although this is not the case of the bulk conduction band energies.

The two preceding examples show that a TB model which can predict both conduction band energies and derivatives does not yet exist. One would have certainly to take into account an extended basis including s, p and d atomic orbitals.

FIT OF EMPIRICAL TB PARAMETERS

Due to the wave vector conservation law, the electron-hole radiative recombination yield in bulk silicon which is an indirect gap semiconductor is very small as it must be phonon assisted. It is improved for silicon clusters as this selection rule is no more valid. There is some interest to consider the BC-8 phase which is a direct gap semiconductor [24]. For this BC-8 phase, to get the same quality of the fit for the band structure as for the diamond structure one has to consider TB interactions up to 0.5 nm. This corresponds to the 7th nearest neighbors in the BC-8 phase which is close to the 3rd neighbors in the diamond phase. Together with the atomic levels E_s and E_p, there are 58 TB parameters. To avoid a too large number of parameters whose numerical fit might lead to unphysical values, we have expressed the s and p atomic functions as linear combination of two gaussians and we have fitted the amplitude A and the exponential coefficient α of each gaussian. Two-centers analytical expressions for the four overlap parameters S_i (i=ssσ, spσ, ppσ and ppπ) between gaussian orbitals as a function of the interatomic distance are given in reference [25]. Four other extra prefactors E_i are used for the hamiltonian parameters H_i which are written as E_iS_i but with α and A parameters different from the overlap ones. We are left with 28 parameters which have been fitted to a band structure obtained by norm conserving pseudopotential theory [26] within the local density approximation.

The fitting procedure can be simplified if one uses group theory to block diagonalize the hamiltonian as it is often not sufficient to sort the eigenvalues by increasing values because their order can change during the fit. When there are several atoms in the unit cell (8 in the case of the BC-8 structure,) this analytic calculation becomes rapidly tedious. This can be done numerically by a simple trick which is lattice structure and unit cell independent.

Let us first take rough estimates of the parameters taken as Harrison's laws [27]. We first compute the n eigenvalues and eigenstates $\phi_{n,k}$ at a symmetry k point. The states $\phi_{n,k}$ have the symmetry of the unit cell. If now we take random TB parameters, the hamiltonian matrix in this basis is no more diagonal but can be block-diagonalized. In the basis of the $\phi_{n,k}$ states, the hamiltonian matrix is for example:

$$H = \begin{bmatrix} x & 0 & x & 0 & 0 & 0 & 0 \\ 0 & x & 0 & 0 & 0 & 0 & 0 \\ x & 0 & x & x & 0 & x & 0 \\ 0 & 0 & x & x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x & 0 & x \\ 0 & 0 & x & 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 & x & 0 & x \end{bmatrix} \qquad (3)$$

where x stands for a non-zero matrix element. The zeroes are due to the symmetry of the $\phi_{n,k}$ states. Looking for the non-zero matrix elements in this matrix, it is possible to reorder the $\phi_{n,k}$ states. In the example given in equation (3), we get a non-zero 4*4 block for $n=1, 3, 4$ and 6, a 1*1 for $n=2$ and a 2*2 for $n=5$ and 7. If we reorder the functions $\phi_{n,k}$, we get

$$H = \begin{bmatrix} x & x & 0 & 0 & 0 & 0 & 0 \\ x & x & x & x & 0 & 0 & 0 \\ 0 & x & x & x & 0 & 0 & 0 \\ 0 & x & x & x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x \end{bmatrix} \qquad (4)$$

which is now block diagonalized. If a group representation of dimension m is l times degenerated, this procedure generally gives a $lm*lm$ block because the $\phi_{n,k}$ computed by a diagonalization program are a random combination of the true functions which give l identical $m*m$ blocks. This is not a real problem for the m eigenvalues of the $lm*lm$ block are l times degenerated and can be easily related to the degenerated first principles energies.

The values we have calculated (Fig. 7) are close to the results obtained for silicon in the diamond structure [11] and they give silicon BC-8 conduction band (Fig. 8) in very good agreement with the LDA result. The energy gaps computed for Si BC-8 clusters are given on Fig. 9. The agreement with a LDA calculation for small clusters is quite good.

Let us also notice two points which can improve the quality of the fits. The variations of the TB parameters with interatomic distance are sometimes computed by a fit using several first principles band structures calculated for different lattice structures. But these band structures have no common energy reference. The zero energy is usually taken at the Fermi level or at the top of the valence band. So the hamiltonian parameters have to be written as $H_i-\Delta E_\beta*S_i$ where ΔE_β is an extra parameter which depends on the lattice structure β. Such an expression must also be used in the case of a fit to a single band structure as in this case changing the origin of energy would modify the quality of the fit.

The second point concerns the values of the TB parameters used as the starting point of the fit. As we have to solve a non-linear system of equations to minimize the error equal to the distance between the TB eigenstates and the first principles ones, one can finds different minima depending on the starting point. The errors obtained for these different minima are generally very close and it
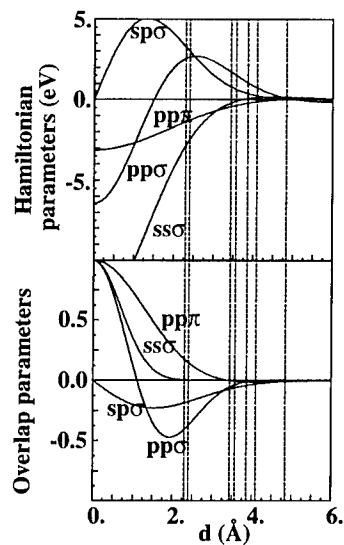
Fig. 7 Non orthogonal tight-binding parameters for Silicon in the BC8 phase. The vertical dotted lines indicate the positions of the nearest neighbors.
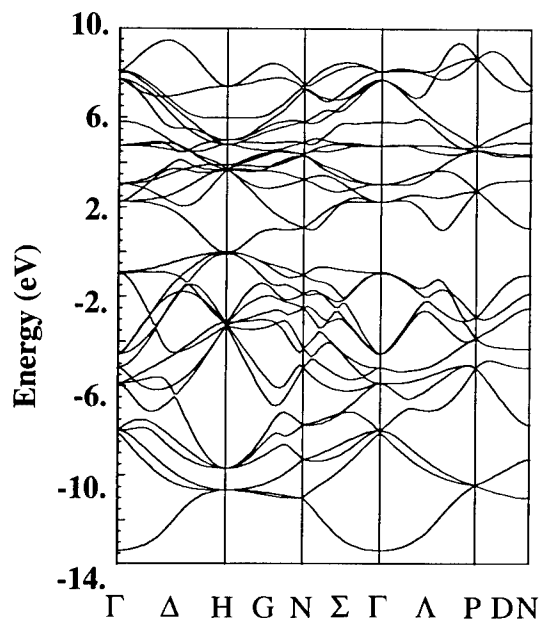
Fig. 8 .Tight-binding band structure of BC8 silicon
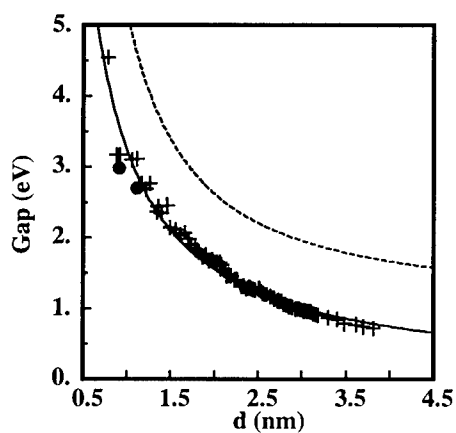
Fig. 9 Variation of the cluster band gap as a function of the cluster diameter for silicon in the diamond structure (dotted line) and in the BC8 phase: tight-binding model (crosses); LDA calculation (full dots). The full line is a mean least square fit to the calculated points
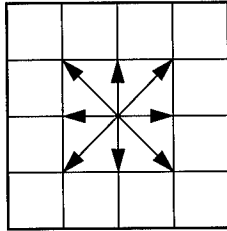
Fig. 10 In a square lattice and TB interactions limited to 2nd nearest neighbors an atom of a 4*4 unit cell has no interaction with an atom of a neighboring cell.

is difficult to choose between the corresponding TB parameters sets. To find a starting points or the TB parameters without any fit, let us write the hamiltonian in a basis of local atomic functions. The basis generally contains many states from the neutral atoms and the ionized ones. We also choose a unit cell large enough such as an atom will have no TB interaction with an atom in a neighboring cell (Fig. 10). We first orthogonalize the basis and write the hamiltonian as:

$$H = \begin{bmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{bmatrix}$$ (5)

where A is a minimal basis set (sp$_3$ for a silicon atom for example) and B the remaining basis functions higher in energy. To keep the size of the problem equal to the dimension of A, we have to consider an effective Hamiltonian:

$$H_{eff} = H_{AA} + H_{AB}\frac{1}{E - H_{BB}}H_{BA}$$ (6)

The matrix elements of H_{eff} give the TB parameters. If their variation with energy E in the A states spectrum is not negligible, one can conclude that the minimal basis set A is too small. On the contrary, to simplify the search of the H_{eff} eigenstates a constant E can be taken.

Such method can be also used to get the variation of the intra-atomic TB energy levels and of the two- and three-centers interatomic integrals when the interatomic distance or the atom environment are modified.

CONCLUSION

We have made a pretext of the calculation of silicon clusters electronic structure for showing how TB can be used to get accurate results for sizes where neither EMA nor first principles can be applied. It also shows that to get accurate cluster band gaps, silicon is certainly among the most difficult materials. This is due to the difficulty of accurately describing the conduction band. This is certainly due to the d states generally omitted in silicon TB calculations. Some improvements have been made to take into account the symmetry to simplify the fit of the TB parameters. A method to avoid the fit of the TB parameters is also suggested. It can also be used to test the validity of the minimal basis set used in TB calculations.

ACKNOWLEDGMENTS

Centre National de la Recherche Scientifique".

REFERENCES

1 L.T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
2 P. Vogl, H.P. Hjalmarson, and J. Dow, J.Phys.Chem.Solids **44**, 365 (1983).
3 M. Lannoo and G. Allan, J. Phys. Chem. Solids, **32**, 637 (1971).
4 A. Zunger, J. Phys. C **7**,76 (1974).
5 R. P. Messmer, Phys. Rev. B **15**, 1811 (1977).
6 O. Bilek and L. Skala, Czech. J. Phys. B **28**, 1003 (1978).
7 P. Kadura and L. Kunne, Phys. Status Solidi B **88**, 537 (1978).
8 R. A. Smith, *Wave Mechanics of Crystalline Solids* (Wiley, New York, 1961), p. 62.
9 M. V. Ramakrishna and R.A. Friesner, Phys. Rev. Lett. **67**, 629 (1991); J. Chem. Phys. **95**, 8309 (1991); **96**,873 (1992).
10 S.H. Zhang, C.-Y. Yeh, and A. Zunger, Phys. Rev. B 48, 11204 (1993).
11 L.F. Mattheiss and J.R. Patel, Phys.Rev.B **23**, 5384 (1981); P.B. Allen, J.Q. Broughton and A.K. McMahan, Phys.Rev.B **34**, 859 (1986).
12 E. Kauffer, P. Pecheur, M. Gerl J.Phys. **C9**, 2319 (1976).
13 C. Tserbak, H.M. Polatoglou, and G. Theodorou, Phys. Rev. B **47**, 7104 (1993).
14 J.P. Proot, C. Delerue and G. Allan, Appl.Phys.Lett. **61**, 1948 (1992); C. Delerue, G. Allan and M. Lannoo, Phys.Rev.B **48**, 11024 (1993).
15 C. Delerue, G. Allan, and M. Lannoo, *"Light Emission in Solids: From Physics to Devices"*, Lockwood Ed., "Semiconductors and Semimetals" Vol. 49, Academic Press (London - 1998).
16 B. Delley and E.F. Steigmeier, Appl.Phys.Lett. **67**, 2370 (1995).
17 D.J. Lockwood, Solid State Communications **92**, 101 (1994).
18 M. Ehbrecht, B. Kohn, F. Huisken, M. A. Laguna, and V. Paillard, Phys. Rev. B **56**, 6958 (1997).
19 F. Koch, V. Petrova-Koch and T. Muschik, Journal of Luminescence **57**, 271 (1993).
20 G. Allan, C. Delerue, and M. Lannoo, Phys.Rev.Lett. **76**, 2961 (1996).
21 G. Allan (Unpublished results).
22 G. Allan, C. Delerue, and M. Lannoo, *"Advances in Microscrystalline and Nanocrystalline Semiconductors - 1996"*, P. M. Fauchet, I. Shimizu, R. W. Collins, P. A. Alivisatos, J.-C. Vial, T. Shimada Eds (Mat. Res. Soc. Proc. **452**, Pittsburgh, PA, 1997) p. 75
23 DMol User Guide, version 2.3.5. San Diego: Biosym Technologies, 1993.
24 J. D. Joannopoulos, and M. L. Cohen, Phys. Rev. B **7**, 2644 (1973)
25 R. C. Chaney, T. K. Tung, C. C. Lin, and E. E. Lafon, J. Chem. Phys. **52**, 361 (1970)
26 N. Troullier, and J.L. Martins, Solid State Comm. **74**, 613 (1990).
27 W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

STRUCTURE, BONDING AND STABILITY OF TRANSITION METAL SILICIDES: A REAL-SPACE PERSPECTIVE BY TIGHT BINDING POTENTIALS

Leo MIGLIO, Francesca TAVAZZA, Antonio GARBELLI and Massimo CELINO°
Istituto Nazionale di Fisica della Materia and Università di Milano,
Dipartimento di Scienza dei Materiali, 15 via Emanueli, I-20126 MILANO, Italy.
°Ente per le Nuove Tecnologie, ENEA, CRE della Casaccia, PO 2400, 00100 Roma, Italy

ABSTRACT

We point out that the predictive power of tight binding potentials is not limited to obtaining fairly accurate total energy calculations and very satisfactory structural evolutions by molecular dynamics simulations. They also allow for a nice physical picture of the links between bonding and stability in different structures, which is particularly helpful in the case of binary silicides.

INTRODUCTION

Recently, semi-empirical tight binding (TB) potentials have been successfully adopted for total-energy calculations and numerical simulations of semiconductors, transition metals and carbon-based materials. An increasing amount of applications are taking place whenever the complexity or the size of the system prevents the employment of first-principles approaches, and We think that this volume contains the most complete and updated review on the subject. Probably, one reason for such an accomplishment is related to the low computational cost (as compared to selfconsistent methods) entailed by the semiempirical estimation of the electronic states entering the attractive (band structure) part of the potential (U_{bs}), and by the phenomenological description of the repulsive part (U_{rep}) through pair interactions. One other reason is represented by the very recent achievements in developing linear scaling algorithms that may encompass the bottleneck of this non-selfconsistent method, i.e. the TB matrix diagonalization. They have disclosed the possibility of large scale simulations for realistic materials and a complete, updated review can be found in [1].

Still, the most interesting feature of this method, in our opinion, rests in the possibility of exploiting a real-space analysis of the electronic features, and in the interpretative power provided by the partition of the total energy (and interatomic forces) into band structure and repulsive contributions. For what concerns the former, the orbital- (α) and site- (i) projected density of states $D_{i\alpha,i\alpha}(\varepsilon)$ can be very helpful, especially in the case of non-periodic systems, where the wavevector integration on k is reduced to the Γ point of a large supercell. Moreover, the evaluation of the crystal orbital overlap populations (COOP)

$$D_{i\alpha,j\beta} (\varepsilon) = \text{Re}[\Sigma_{nk}C^*_{i\alpha}(nk)C_{j\beta}(nk) \ S_{i\alpha,j\beta} \ \delta(\varepsilon-\varepsilon_{nk})]$$

can be achieved also in the case of orthogonal tight binding by substitution of the overlap element $S_{i\alpha,j\beta}$ with the corresponding hopping element $H_{i\alpha,j\beta}$, in agreement to the Hückel guess $S_{i\alpha,j\beta} = 2H_{i\alpha,j\beta}/K(\varepsilon_{i\alpha}+\varepsilon_{j\beta})$. This allows for a qualitative analysis of bonding and antibonding features in the density of states (DOS) and for the evaluation of the bond energy contributions to U_{bs}

$$U_{i\alpha,j\beta}(\varepsilon_F) = \int D_{i\alpha,j\beta}(\varepsilon) \ d\varepsilon.$$

Obviously, this is not a novelty in the standard LCAO analysis of the stability trends in solids, as nicely outlined in the last fifteen years by distinguished chemists like Roald Hoffman [2], or physicists like David Pettifor [3]. What makes still vital the subject is the need to import it into the standard analysis of molecular dynamics simulations, which is still mainly grounded on structural informations, such as the ones provided by pair correlation distributions.

The repulsive potential, on the other hand, also convey some additional information, if we consider that (within a simple pair potential scheme) it contains the orbital overlap correction and the crystal field contribution. In fact, competing structures may display a different partition in U_{bs} and U_{rep}, depending on the neighbours configurations, in terms of covalent bonding and Pauli repulsion, respectively. Clearly, in case of binary compounds, the need to extend interactions up to secon neighbours (as motivated by the achievement of a satisfactory DOS) may introduce some arbitrariness in the determination of the pair repulsion parameters. Still, a wise fitting on the equilibrium and stability conditions for a few structures always gives rise to a phenomenological link between the atomic size and the repulsive character.

In the last few years we have applied the orthogonal TB method to total energy calculations and molecular dynamics simulations of Ni-, Co- and Fe- silicides [4-12]. The silicon-rich compounds are particularly interesting from this point of view, since they display both a high sensitivity of the electronic features to the bond directions, as provided by the fairly covalent p_{Si}-d_{TM} bonding, and a relevant polymorphic attitude, which is typical of metallic materials. The latter feature is emphasized in the case of thin epitaxial films, which frequently display pseudomorphic (i.e. substrate-induced) crystal phases, not present in the bulk phase diagram. One example is $FeSi_2$, which is presently the subject of a vivacious interst, due to the prosimising performances provided by its stable, orthorombic phase as a light emitting diode integrated at a silicon junction [13].

ONE EXAMPLE: $FeSi_2$

At variance with respect to the related compounds $NiSi_2$ and $CoSi_2$, the metallic fluorite phase γ-$FeSi_2$ is not bulk-stable, and evolves towards the the β form, which exhibits a semiconductive gap as large as 0.8-0.9 eV [14]. The latter includes 24 atoms in a primitive, base-centered orthorhombic cell, and it is generated by a moderate deformation of the cubic cages of silicon atoms around the iron sites in the fluorite structure, leading to inequivalent Fe_I, Fe_{II}, Si_I and Si_{II} sites, with sizeable changes in the secondary coordination and volume reduction by 2.9 % [15]. Actually, muffin tin orbitals calculations by Christensen [16] have predicted a high density of states at the Fermi level of the γ-phase, suggesting that a Jahn-Teller distortion drives the cubic structure into the β form. Some more insight is provided by our analysis of the COOP for p_{Si}-d_{Fe} bonding (see Fig.1, top panel), within the TB potential that we have developed for iron silicides [10]. It shows that the Fermi level is located at the maximum of an antibonding (i.e. negative) $D_{i\alpha,j\beta}$ in the case of the fluorite structure. On the contrary, the lattice distortion provided by the β phase gives rise to an enhancement of the bonding features at low energy (nearly the double of the bond energy) and to a splitting of the antibonding band where the Fermi level is located (see Fig. 1, bottom panel).

However, recent Molecular Beam Epitaxy experiments have shown that the γ phase can be stabilized at very low coverages ontop Si (111) due to the bad matching of the β phase to the substrate and to the superior interface bonding provided by the fluorite arrangement [17]. As the film exceeds a few monolayers, a real structural phase transition to the stable form occurs at annealing temperature which decreases with thickness. A similar behaviour is also found for other pseudomorphic phases of $FeSi_2$, such as the CsCl-defected (d) structure [9] and the α-phase
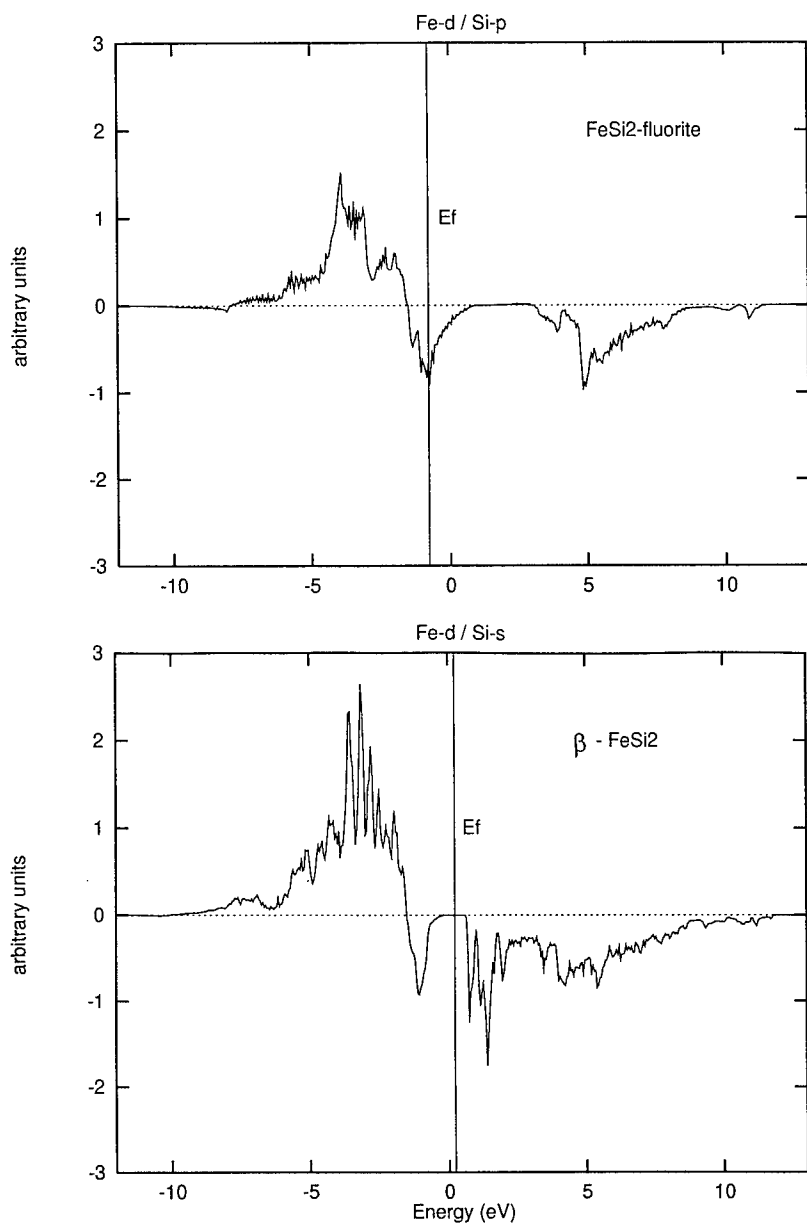
Fig. 1. COOP estimation by our TB potential for p_{Si}-d_{Fe} first neighbours bonding in fluorite (top panel) and in orthorhombic (bottom panel) structures.

(bulk stable at high temperatures), which are closer in total energy to the stable phase (see our TB total energy estimation in Fig. 2, as taken from [6]), still displaying large kinetic barriers to the selected atomic jumps within a pseudocubic network that the corresponding phase transitions do require.
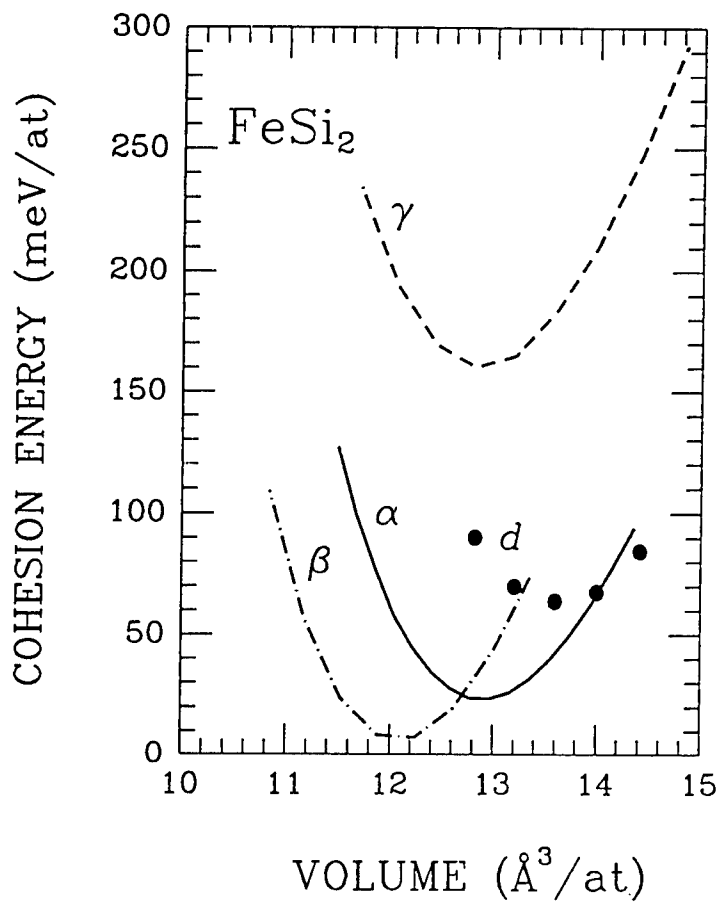
Fig. 2. Total energy per atom vs atomic volume for the existing FeSi$_2$ phases, as calculated by the TB potential described in [6].

Here, by Tight Binding Molecular Dynamics (TBMD), we are interested in analyzing from a site-projected point of view the spontaneous evolution of the unstable bulk fluorite structure towards the stable phase, in order to see whether it corresponds to a Jahn-Teller mechanism. One may wonder how much our bulk simulation at 100 K does reproduce the structural phase transition occurring in the epitaxial configuration. Since we are not interested in the thermodynamic aspects, and since we show that the driving force of the transformation lies in the coupling between the local electronic density of stated and the local lattice distortion, which is obviously hindered by the lattice-matching condition at the silicon interface, we conclude that the only difference with respect to the epitaxial configuration rests in the "extrinsic" kinetic barrier provided by the interface bond breaking.

Elsewhere [10] we have described how we implemented a suitable Rahman-Parrinello molecular dynamics with variable cell shape, and how we selected a suitable cutoff of the TB and repulsive interactions, in order to include the second neighbours shell. Here we focus only on the main physical results that we obtained during a simulation as long as 10^5 time steps ($\tau=0.2$ fs), which can be analyzed within the real space perspective that we outlined above. The simulation cell is initially taken as the tetragonal nucleus (x-oriented) of a cubic supercell made by 8 fcc units, where a is two times the fcc edge (5.37 Å) and b,c are the diagonals of the fcc square faces. It contains 48 atoms and, as the simulation evolves, it transforms into a simple orthorhombic cell (containing two primitive units), by decreasing a and increasing b and c very close to the experimental values values (a=9.86 Å, b=7.79 Å, c=7.83 Å). Even the final atomic positions inside the simulation cell agree very well to the experimental target structure, as confirmed by the standard analysis of the radial and bond-angle pair distribution functions [10] and the site coordinations reported in Fig.3.

In Fig.4 we follow the evolution of the electronic DOS with simulation time, which shows that a metal-semiconductor transition takes place. It is interesting to note that, after the initial disappearance of the sharp density peak at the Fermi energy (0-200 τ) as a consequence of the thermal symmetry breaking, a spread out of low energy s states occurs inbetween 200 and 3000 τ, along with a density redistribution of p/d states above and below the Fermi energy. Still, the gap opening takes place only between 4150 and 4700 τ and the DOS does not change any further, showing the same shape and gap size of the *ab intio* result reported in [14].

In order to tell if a true Jahn-Teller distortion is taking place, we now analyze some more in details the configurational path during the first 5000 τ (= 1 ps, the intrinsic relaxation time of our simulation cell). In Fig. 5 we display the evolution of the secondary coordinations for Fe and Si sites, Fe-Fe and Si-Si respectively, since the primary coordination for Si-Fe pairs is always eight, despite the relevant bond angle distortions (see Fig. 3). We note that between 2000 and 2700 τ a sizeable change occurs and that the transformation is completed in a second, separate step between 4000 and 4700 τ. It would be interesting to understand whether a partial, cooperative motion at all the sites is involved in both steps, or if the two-step behaviour is produced by a complete, local transformation occurring at different sites in different times.

To this end we show in Fig. 6 the site projected DOS at 3000 τ, the intermediate stage, for a selected number of atoms representing 6 different classes in which the atoms in the cell turn out to be divided. Four Fe_I atoms over eight, MI (A), do display a gap and the local environment (bond angles and secondary coordination) is entirely of β-type. The remaining four, MI (B), are essentially fluorite-like and no local gap is present. The eight Fe_{II} sites, MII (AB, BA) are in an intermediate stage both for what concerns the DOS and the structure, and they stay in this situation up to the second structural rearrangement at 4000 τ, when the remaining Fe_I sites also make the transition.

313
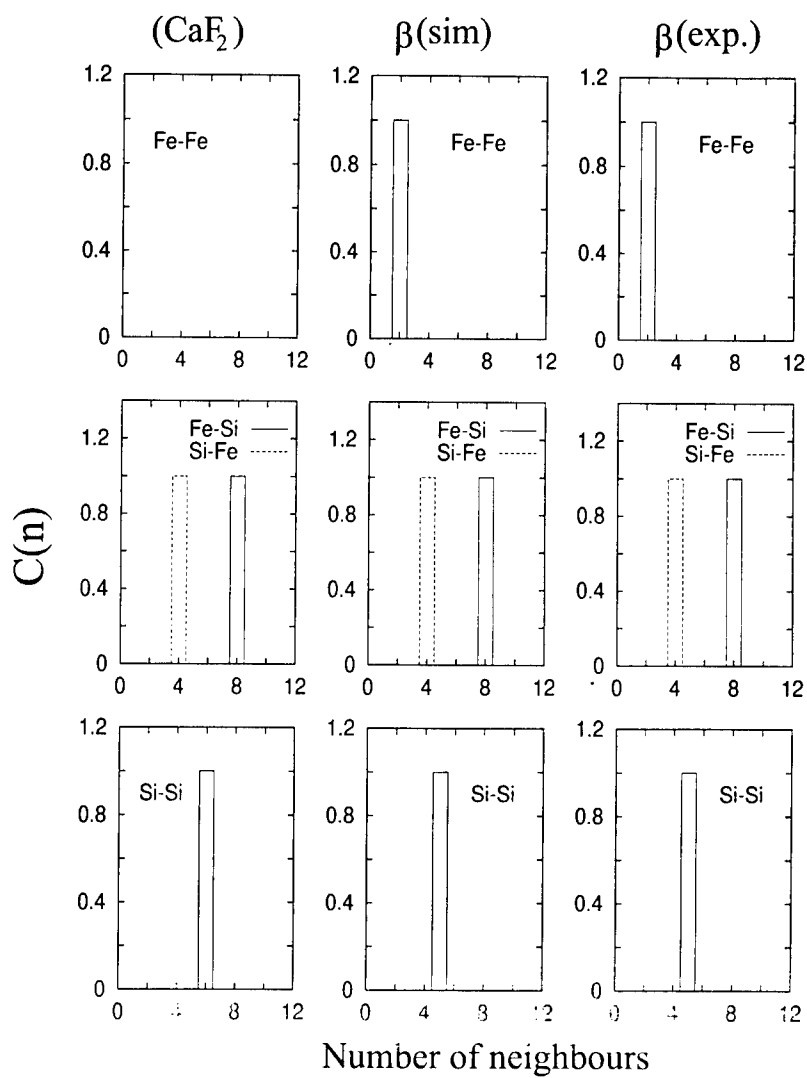
Fig. 3. *Site coordination distribution functions for second (Fe-Fe, Si-Si) and first (Fe-Si) neighbours in the fluorite (left column), experimental orthorhombic (right column) and simulated orthorhombic (center column) configurations.*
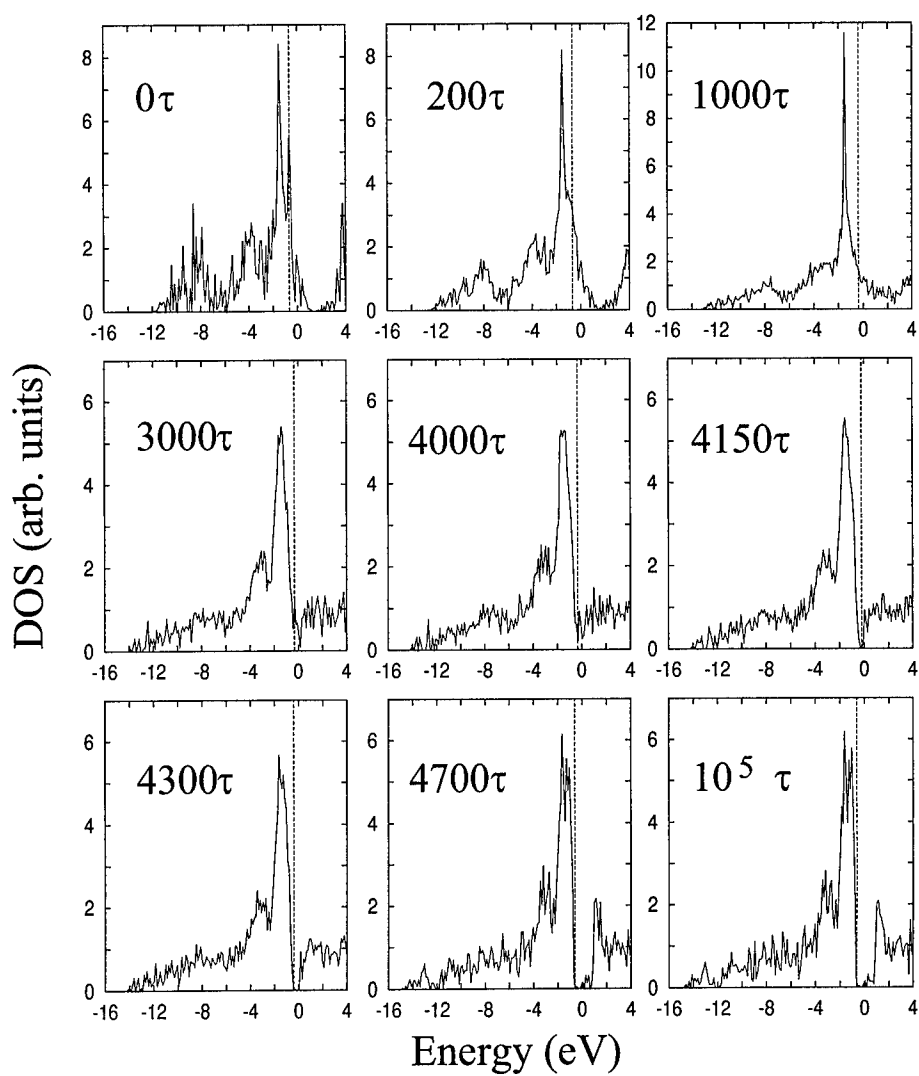
Fig. 4. Evolution of the electronic DOS with simulation time. The actual Fermi level position is indicated by a vertical dashed line.

Both Si_I and Si_{II} sites, 16 atoms in each class, are sharply divided into two groups of eight atoms: one, SiI,II (A), contains completely transited environments (the distorted cages around Fe_I transited atoms) with a local gap opening, and the remaining, SiI,II (B), are still fluorite-type. Therefore a set of local distortions take place in the first step, and they are distributed in an ordered pattern indipendently of the final Fe_I or Fe_{II} symmetry.
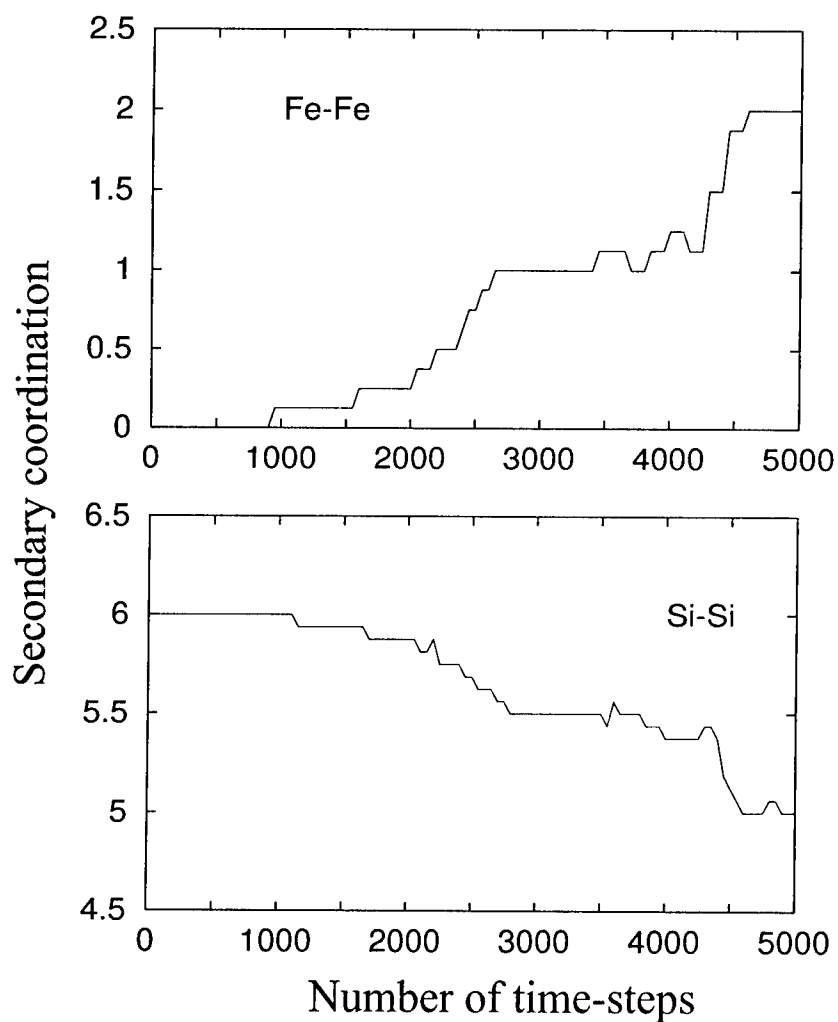
Fig. 5. Evolution of the secondary coordination for Fe (top panel) and Si (bottom panel) sites.

DOS (arb.units)
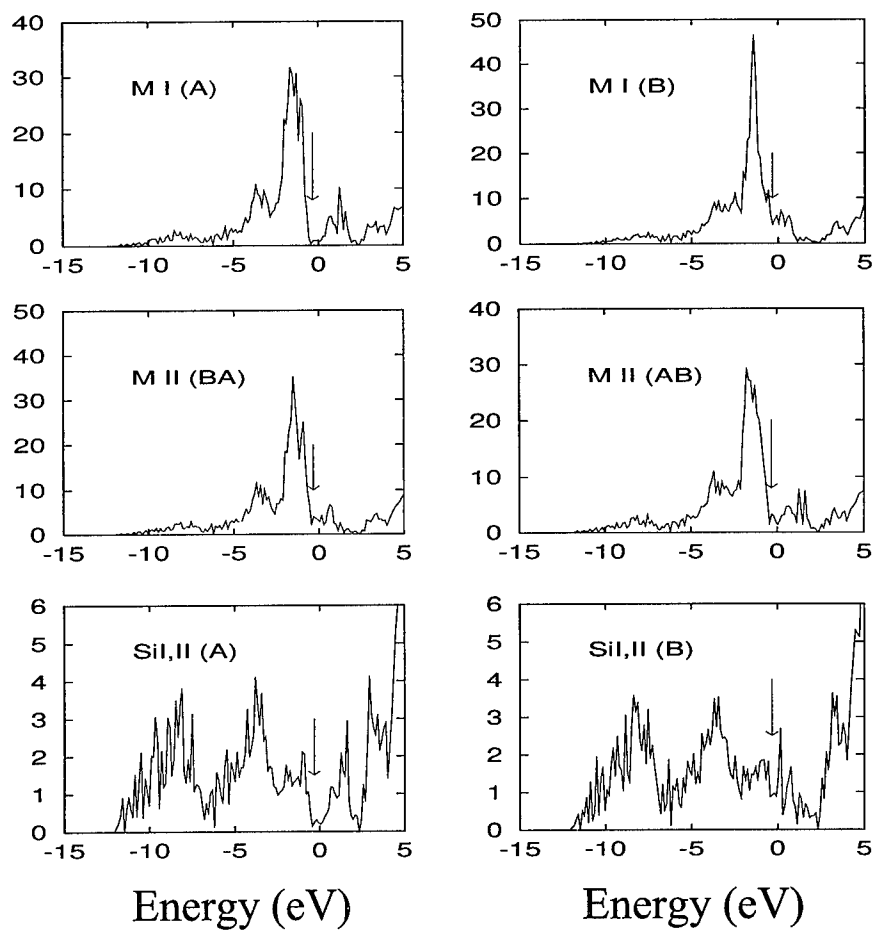
Fig. 6. Site projected DOS at 3000 τ for metal (top four panels) and silicon (two bottom panels) sites, as classified according to the text.

In Fig. 7 we show the comparison between the the original cubic arrangement in fluorite and the distorted cage which gives rise to a local gap opening, as extracted by our simulation cell at 3000 τ. We see that a displacement of Fe (small sphere) towards one face of the cubic cage occurs, generating a bond angle distortion at first neighbours, which is responsible of the gap appearance, according to the rather covalent character of the Fe_d-Si_p bonding. This, in turn, gives rise to a deformation of the silicon cage into a solid with two square faces of different side and four nearly regular trapezoidal faces.
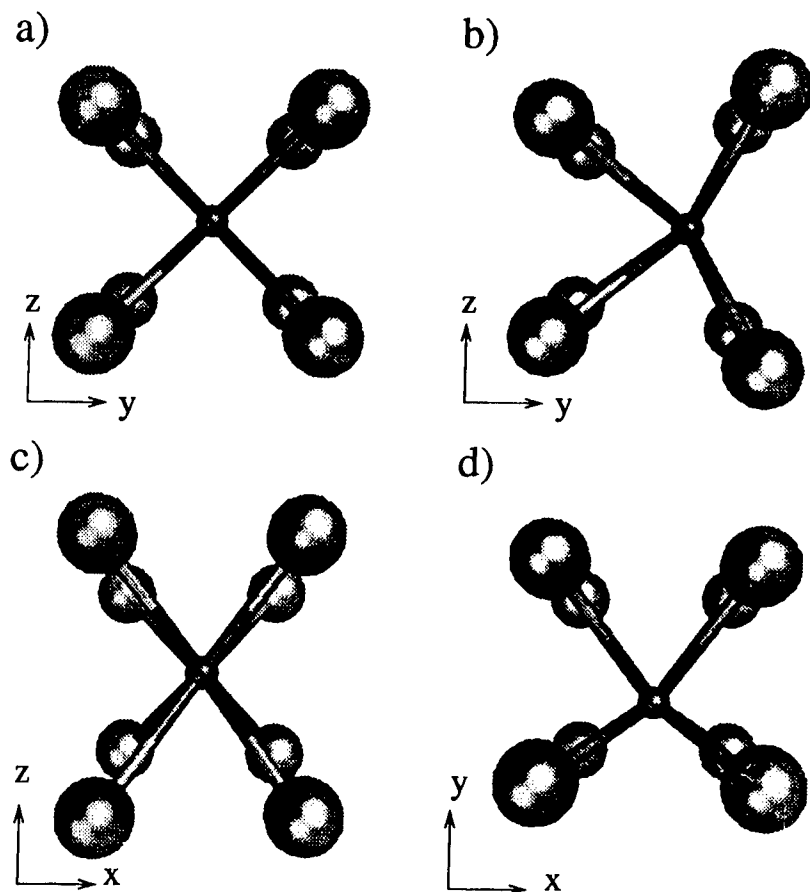
Fig. 7. *Prospective views of fluorite (a and c panels) and orthorhombic (b and d panels) local configurations, as taken along x and z axis, respectively. Small spheres are the iron atoms.*

One additional information is provided by the analysis of the repulsive potential for the two structures at the same volume per atom. The β configuration, with respect to the γ, turns out to have not only a larger U_{bs} , but also a lower U_{rep}, as probably due to the slight increase in Fe-Si first neighbours distance that the silicon cage distortion generates. This occurs despite the sizeable reduction in the second neighbours distances, which are still within the cutoff radius of our potential. By the analysis of U_{rep} at different steps of our simulation, we understand that the absence of any activation barrier can be related to the fact that no increase in the Pauli overlap is generated during the transformation, and in the target structure too. This is not the case, for example, of the defected (d) and α pseudomorphic structures.

CONCLUSIONS

In this paper we have just considered some qualitative aspects concerning one example of structural transformation in transition metal silicides, which can be analyzed by TBMD. Quantitative details on the subject (but for the analysis of the COOP and the repulsive potential) are contained in [10]. The important point that we hope to have made sufficiently clear, however, is that the TB method provides more information that is usually extracted from molecular dynamics simulations, especially if we take profit of the real-space analysis of the bonding and the repulsive features.

REFERENCES

1. G. Galli, Current Opinion in Solid State & Mat. Sci. **6**, 864 (1996).

2. R. Hoffmann, Solids and Surfaces: A Chemist's view of Bonding in Extended Structures, VCH Publishers Inc., New York, 1988.

3. D. Pettifor, Bonding and Structure of Molecules and Solids, Claredon Press, Oxford, 1995.

4. G. Malegori and L. Miglio, Phys. Rev. B **48**, 9223 (1993); H. von Kaenel, M. Mendik, K.A. Maeder, N. Onda, S. Goncalves-Conto, and C. Schwarz, G. Malegori and L. Miglio, F. Marabelli, Phys. Rev. B **50**, 3570 (1994).

5. Leo Miglio and Giovanna Malegori, Phys.Rev. B **52**, 1448 (1995).

6. Leo Miglio, Francesca Tavazza and Giovanna Malegori, Appl. Phys. Lett. **67**, 2293 (1995).

7.H. von Kaenel, C. Schwarz, S. Goncalves-Conto, E. Mueller, L. Miglio, F. Tavazza and G. Malegori, Phys.Rev.Lett., **74**, 1163 (1995).

8. S. Sanguinetti, C. Calegari, V.R. Velasco, G. Benedek, F. Tavazza and Leo Miglio, Physical Review B **54**, 9196 (1996).

9. S. Goncalves-Conto, U. Schaerer, E. Mueller and H. von Kaenel, L. Miglio and F. Tavazza, Phys. Rev. B **55**, 7213 (1997).

10. Leo Miglio, Valeria Meregalli, Francesca Tavazza and Massimo Celino, Europhys Lett. **37**, 415 (1997); Francesca Tavazza, Valeria Meregalli, Leo Miglio and Massimo Celino, submitted to Phys. Rev. B.

11. G. Guizzetti, F. Marabelli, M. Patrini, P. Pellegrino, and B. Pivac, Leo Miglio and V. Meregalli, H. Lange, W. Henrion, V. Tomm, Phys. Rev. B **55**, 14290 (1997)

12. M. Fanciulli, G. Weyer and J. Chevallier, H. von Kaenel, H. Deller and N. Onda, L. Miglio, F. Tavazza and M. Celino, Europhys. Lett. **37**, 139 (1997).

13. D. Leong, M. Harry, K.J. Reeson, and K.P. Homewood, Nature **387**, 686 (1997).

14. R. Eppenga, J. Appl. Phys. **68**, 3027(1990); S. Eisebitt, J.E. Rubensson, M. Nicodemus, T. Böske, S. Blügel, W. Eberhardt, K. Radermacher, S. Mantl, and G. Bihlmeyer, Phys. Rev. B **50**, 18330 (1994); J. van Ek and P.E.A. Turchi, P.A. Sterne, Phys. Rev. B **54**, 7897 (1996); A.B. Filonov, D.B. Migas, V.L. Shaposhnikov, N.N. Dorozhkin, G.V. Petrov, V.E. Borisenko, W. Henrion, and H. Lange J. Appl. Phys. **79**, 7708 (1996).

15. Y. Dusausoy, J. Protas, R. Wandji, and B. Roques, Acta Cristallogr. Sect. B **27**, 1209 (1971).

16. N.E. Christensen, Phys. Rev. B **42**, 7148 (1990).

17. H. von Kaenel, R.Stalder, H.Sirringhaus, N.Onda , J.Henz, Appl. Surf. Sci. **53**, 196 (1991); A.L.Vazquez de Parga, J.De la Figuera, C.Ocal R. Miranda, Europhys. Lett. **18**, 595 (1992); Le Thanh Vinh, J.Chevrier J.Derrien, Phys. Rev. B **46**, 15946 (1992); H.Ch.Schaefer, B.Roesen, H.Moritz, A.Rizzi, B.Lengeler, H.Luth, D.Gerthsen, Appl. Phys. Lett. **62**, 2271 (1993). A. Rizzi, B. N.E. Rösen, D. Freundt, Ch. Dieker, H. Lüth, and D.Gerthsen, Phys. Rev. B **51**, 17780 (1995).

DEVELOPMENT OF SIMPLE SPD TIGHT-BINDING MODELS FOR TRANSITION METALS

O. Le Bacq, F. Willaime

Section de Recherches de Métallurgie Physique, CEA Saclay, 91191 Gif-sur-Yvette, France

A. Pasturel

Maison des Magistères, EP-96, CNRS, BP 166, 38042 Grenoble, France

ABSTRACT

A simple *spd* tight-binding scheme for atomistic simulations in transition metals is developed using an orthogonal basis set in the two-center approximation. The purpose of the present approach is to cope with the limitations of the canonical *d*-band model for elements at the beginning or at the end of the transition metal series while keeping a reduced number of parameters, and simple decay-functions. The parameters for the hopping integrals are fitted to FP-LMTO calculations of the volume dependence of band energies and band structures for several selected structures. Constant values are taken for the on-site energies, and the Born-Mayer pair potential is used for the repulsive term. Two approaches are compared for the total energy: the *band* model and the *bond* model. First applications are presented in the case of zirconium, where the difference between these models on phase stability results is particularly drastic. The *bond* model reproduces the stability of the hcp phase and displays a good agreement with experiments for the elastic constants.

INTRODUCTION

Large scale atomistic simulations in transition metals increasingly need the use of energetic models being both fast and quantitative. The realism reached by calculations based on the Local Density Approximation (LDA) to the Density Functional Theory (DFT) is very satisfactory in most cases. However their computational cost are quite a limitation to reach very large system sizes, and to perform long runs of molecular dynamics simulations in particular in metals. Even if the size of the matrices which are handled can be reduced by using localized basis sets, the calculation of the matrix elements and of the various contributions to the total energy together with the self-consistent loop make these calculations typically two orders of magnitude heavier or more than tight-binding schemes. For transition metals, *spd* tight-binding models therefore seem to be an interesting way of approaching the realism of *ab-initio* calculations while drastically reducing the computation time.

The present study focuses on Group-IV transition metals (Ti, Zr, Hf) for which an explicit treatment of the electronic structure is required to go beyond semi-empirical models [1] in order to study the still open questions related to point and extended defects. Tight-binding *d*-band models were shown to account for some of the specificities of these elements [2], but the effects due to *s* and *p* electrons and to their hybridization with *d* electrons are also expected to be important and have motivated the development of the present models. These elements are a particular difficult test for tight-binding models because 4 structures - namely hcp, bcc, omega and fcc - are very close in energy [3]. After a brief description of the models and the fitting procedure, this paper illustrates the differences between the *bond* model and the *band* model in the case of zirconium.

TIGHT-BINDING APPROXIMATION

Density functional Theory (DFT), using the Kohn-Sham approximation within LDA gives the well-known following expression for the total energy of a system of interacting electrons in a solid :

$$E[n(\mathbf{r})] = \sum_{\alpha,\mathbf{k}} \omega_{\mathbf{k}} S(\epsilon_\alpha(\mathbf{k}) - E_F)\epsilon_\alpha(\mathbf{k}) + F[n(\mathbf{r})] \tag{1}$$

where $n(\mathbf{r})$ is the electronic density, $\epsilon_\alpha(\mathbf{k})$ are the eigenvalues of the Kohn-Sham hamiltonian for given wave-vectors $\mathbf{k}$, E_F the Fermi level, $\omega_{\mathbf{k}}$ the weight of the $\mathbf{k}$-point in the irreducible wedge of the Brillouin zone, and $S(\epsilon - E_F)$ a step-function limiting the summation inside the Fermi surface. This natural separation between the two contributions to $E[n(\mathbf{r})]$ can be mimicked by replacing the eigenvalue sum (or band energy) by expressions derived within tight-binding methods and finding empirical potentials for $F[n(\mathbf{r})]$ [4]. In contrast to this standard *band* formalism, Foulkes and Haydock [5] have suggested setting the zero of the cohesive energy equal to the energy of free atoms and rewriting the total energy as a sum of a bond energy and a pairwise repulsive contribution. If the later is taken to have a simple exponential decay, the total energy of a system in this *bond* formalism [6] takes the following form:

$$E_{tot}^{bond} = \sum_{\alpha,\mathbf{k}} \omega_{\mathbf{k}} S[\epsilon_\alpha(\mathbf{k}) - E_F] \epsilon_\alpha(\mathbf{k}) - \sum_{i,\ell} N_{i\ell} \epsilon_{i\ell}^0 + A \sum_i \sum_{j\neq i} e^{-p R_{ij}} \tag{2}$$

where $N_{i\ell}$ is the number of electrons of each angular momentum ℓ on site i, and $\epsilon_{i\ell}^0$ the ℓth energy level of the free atom. In the present approach, the tight-binding eigenvalues, $\epsilon_\alpha(\mathbf{k})$, are obtained from a Slater-Koster hamiltonian, written in an orthogonal basis set using s, p and d orbitals, in the two-center approximation.

HAMILTONIAN PARAMETERS

The first step of the fitting procedure is the same for both the *band* model and the *bond* model, and consists of retrieving information on the volume dependence of band structure and band energy from $ab - initio$ calculations. Every independent hopping integral and its variation with distance depend on two parameters through:

$$h_\nu(R_{ij}) = h_\nu^0 exp(-q_\nu R_{ij}) \text{ when } R_{ij} \leq R_c \tag{3}$$

where ν stands for the type of interaction ($ss\sigma$, $dd\sigma$, $dd\pi$, etc), and R_{ij} is the distance between the atoms. A smooth decay to zero is ensured by replacing the exponential by a polynomial from R_c to $R_0 = \sqrt{2}R_c$. The polynomials are taken of order five to make first and second derivatives also continuous at R_c and R_0. Constant values are taken for the on-site energy differences: $\epsilon_s^0 - \epsilon_d^0 = K_{sd}$ and $\epsilon_p^0 - \epsilon_d^0 = K_{pd}$. A rigid shift of the bands is allowed by letting ϵ_d^0 free to vary with structure and atomic volume. We will see in one of the next sections that the total energies in the *bond* model is independent of the value of ϵ_d^0 and this remains true for the particular choice adopted here for the *band* model. The hopping integrals and the parameters defining the on-site energies are fitted simultaneously to FP-LMTO calculations of the volume dependence of band energies E_{band}^{LMTO} and energy bands $\epsilon_{\alpha,s}^{LMTO}(\mathbf{k})$ for several selected structures. In other words, we have to minimize the sum,

$$S = \gamma \sum_{\mathbf{k},\alpha} \sum_{s,\Omega} |\epsilon_{\alpha,s\Omega}^{LMTO}(\mathbf{k}) - \epsilon_{\alpha,s\Omega}^{TB}(\mathbf{k})|^2 + \beta \sum_{s,\Omega} |E_{band}^{LMTO}(s,\Omega) - E_{band}^{TB}(s,\Omega)|^2 \tag{4}$$

with respect to fit parameters. We generally use two different structures (s = bcc, fcc) taken for three atomic volumes Ω each. For the band structure part, we have used about two hundred k-points. Integrations for band energy calculations are performed using a uniform mesh in the irreductible wedge of the Brillouin zone. The Hermit-Gaussian smearing method [7] has been employed to approximate the step-function $S(x)$ with a smearing-width $\sigma = 0.3\,eV$. First guess values are obtained by fitting only energy bands (i.e. $\beta=0$). Then the β/γ ratio is progressively increased to improve the band energies. A final value of $\beta/\gamma = 10^5$ is found to give a good compromise in getting accurate band energies while conserving a valuable control on band structures.

RESULTS OF THE FIT AND BAND STRUCTURES

The above procedure has been applied to zirconium, leading to an accuracy of 10^{-3} eV on the six lowest bands selected for the fit. Band structures are very well reproduced in the two structures considered in the fit, namely the bcc structure (see figures 1-(a) and 1-(b)), and the fcc structure (not shown). Figure 1-(c) and 1-(d) compares tight-binding and FP-LMTO energy bands in the hcp phase as a first test of transferability of the parameters.
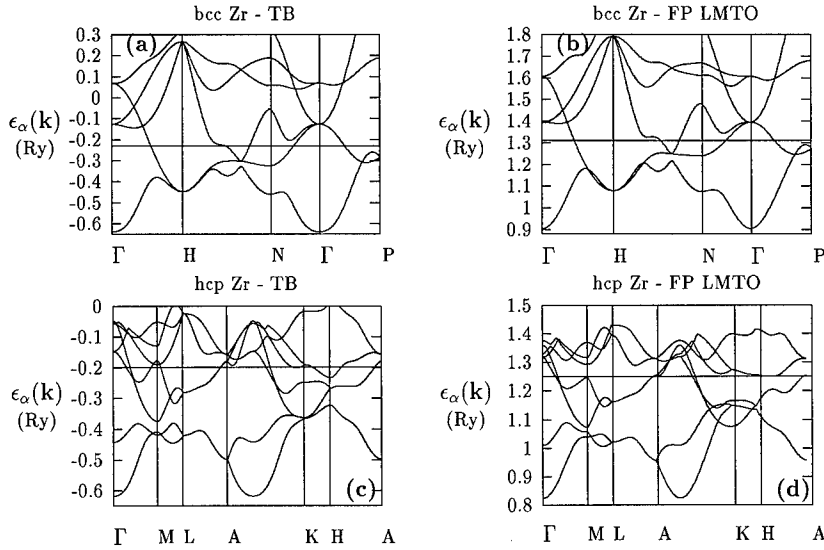
Figure 1: Comparison between spd tight-binding band structures in Zr and the corresponding FP-LMTO bands in (a-b) the bcc phase and (c-d) the hcp phase. $q_{dd\sigma}=0.595$, $q_{dd\pi}=1.167$, $q_{pd\sigma}=0.414$, $q_{pd\pi}=0.311$, $q_{sd\sigma}=0.699$, $q_{pp\sigma}=0.676$, $q_{sp\sigma}=0.648$, $q_{ss\sigma}=0.533$ expressed in $bohr^{-1}$. $h^0_{dd\sigma}=-2.943$, $h^0_{dd\pi}=34.464$, $h^0_{dd\delta}=0.000$, $h^0_{pd\sigma}=-1.208$, $h^0_{pd\pi}=0.160$, $h^0_{sd\sigma}=-4.398$, $h^0_{pp\sigma}=3.781$, $h^0_{pp\pi}=0.000$, $h^0_{sp\sigma}=3.263$, $h^0_{ss\sigma}=-1.364$, $K_{pd}=0.3853$, $K_{sd}=0.1625$ expressed in Ry (1 Ry = 13.6 eV ; 1 bohr = 0.529 Å).

The ability of the model to deal with the hcp phase is quite satisfactory, except for a narrow zone around the K and H points, where discrepancies are noticeable. The parameters characterizing the dd hoping integrals are found to be relatively robust and quite independent of the details of the fitting procedure. It is for instance interesting to notice that the values of the q_ν parameters are very close to those obtained from independent fits performed

on similar models for other transition metals: we may compare $q_{dd\sigma}=1.125$ Å^{-1}, with values of reference [9] for Mo and Re ($q_{dd\sigma} = 1.248$ Å^{-1}) or Nb ($q_{dd\sigma} = 0.987$ Å^{-1}). Moreover the dd parameters were also found to be quite independent from the value of the cut-off radius, R_c. On the other hand, the on-site energy differences change significantly with R_c. The influence of the cutoff-radius, R_c (from 3.491 Å to 6.665 Å), on the band-structure in the hcp-phase, showed that the best transferability was obtained for the smallest value. The hopping integrals are therefore taken as relatively short-ranged, i.e. they do not exceed the third neighbour distance in the bcc phase, in agreement with what is expected from physical arguments on screening effects.

FROM ELECTRONIC STRUCTURE TO ENERGETIC MODELS

The step towards a total-energy model is much more delicate than the band-structure part described above. One way of limiting the empirical character arising from terms other than the hopping integrals and the differences between diagonal elements, is to introduce as few parameters as possible in this second part of the energy. In the *bond* formalism, the bond energy is clearly separated from the repulsive part, which is taken as pairwise here. Moreover, it should be stressed that in a perfect crystal for a pure metal, the total energy in the *bond* model is unambiguously independent from the value of ϵ_d^0, since Eq. (2) can be rewritten as:

$$E_{tot}^{bond} = E_{band}^{(\epsilon_d^0=0)} - N_s K_{sd} - N_p K_{pd} + A \sum_i \sum_{j \neq i} e^{-p R_{ij}}, \qquad (5)$$

where $E_{band}^{(\epsilon_d^0=0)}$ denotes the tight-binding band energy, calculated by setting the on-site energy reference, ϵ_d^0, to zero in the hamiltonian. In the *band* formalism, a constant - like ϵ_d^0 - can be arbitrarily subtracted from the diagonal elements of the tight-binding matrix, if its counterpart is introduced in the second part of the energy [4]. By analogy with d-band calculations, we have chosen to set ϵ_d^0 to zero to calculate the band-energy, and to take the second term in the energy as a pair potential, like in the *bond* model, with a second set of two parameters:

$$E_{tot}^{band} = E_{band}^{(\epsilon_d^0=0)} + A^* \sum_i \sum_{j \neq i} e^{-p^* R_{ij}}. \qquad (6)$$

To summarize, the two formalisms which are compared here have the same number of parameters and are essentially distinguished by the charge transfer between orbitals, when either the volume or the structure is changed. For both formalisms, we have fitted the two parameters of the repulsive potential to the energy-volume curve in the fcc structure, calculated by the FP-LMTO method. The energy-volume curves in the hcp and bcc phases are shown in figure 2. The positive point which is common to both models, is that the different structures have very similar equilibrium volumes and bulk moduli, as desired. The most striking difference is that the relative stability between the hcp and the bcc phases is completely reversed from one model to the other. In the following section we will concentrate on the *bond* formalism which is the only one able to reproduce the stability of the hcp-phase. However, empirical improvements to the *band* model are suggested by analyzing the structure dependence of the values of ϵ_d^0, which are a by product of the band-structure fit. We have indeed noticed that the value of ϵ_d^0 in the bcc structure is higher than in the hcp structure at the same volume by a constant value which is almost independent of the volume. Moreover, if this constant value is taken into account in the *band* model - instead of setting ϵ_d^0 in both structures - the near-degeneracy of the two structures is recovered. However attempts to model this structure dependence of ϵ_d^0 by a simple environment-dependence have failed.

$$E_{tot}^{band} = E_{band}^{(\epsilon_d^0=0)} + E_{rep}$$

$$E_{tot}^{bond} = E_{bond} + E_{rep}$$
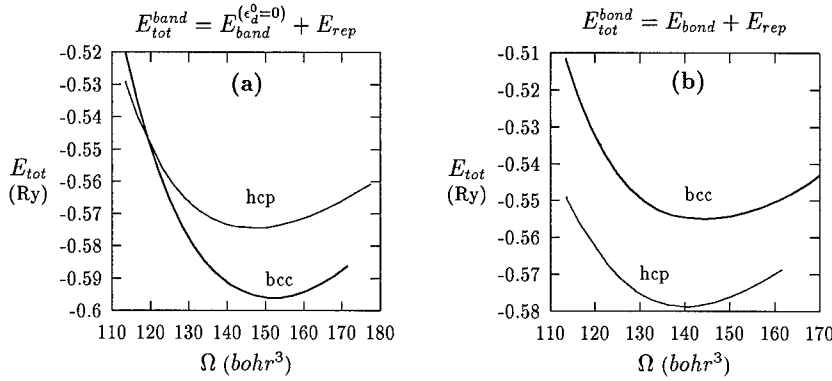
Figure 2: Total energy curves for hcp-Zr and bcc-Zr treated by (a) the band formalism with $\epsilon_d^0 = 0$ (A^*=46.204 Ry and p^*=1.077 $bohr^{-1}$) and (b) the bond formalism (A=13.798 Ry and p=0.817 $bohr^{-1}$).

VALIDATION

In this section, the ability of the present *bond*-model to reproduce the properties of the hcp-phase of zirconium is tested. Since the fits have been performed only on the bcc and fcc phases, the results summarized in Table I have to be viewed as a test of transferability.

The equilibrium lattice parameter is in very good agreement with experiments (1.5 % smaller). The model reproduces the fact that the c/a ratio is smaller than the ideal value of 1.633, but with a discrepancy of 6% with respect to experiments. However this discrepancy is not so large when compared to other quantum mechanical models. The energy difference between compact structures (fcc and hcp) is very satisfactory, but the excess energy of the more open bcc phase is significantly too high when compared to FP-LMTO values.

We have calculated the elastic constants for the ground state hcp-phase of zirconium in order to test the mechanical stability of the crystal and to validate the ability of the model to predict properties connected with second order derivatives of the energy not included in the fit. They are calculated by imposing an external strain on the crystal and performing the calculation of the the second derivative of the total energy with respect to the strain parameter by finite-differences. The energy differences between the strained and unstrained lattice are performed at constant number of electrons. Inner elasticity [8] has been taken into account so that the elastic constants may be written as a sum of two terms, $C_{ij} = C_{ij}^{(0)} + C_{ij}^{rel}$, where $C_{ij}^{(0)}$ is the homogeneous contribution and C_{ij}^{rel} is the relaxation contribution due to relative displacement of the two sublattices of the hcp phase.

	a_0 (Å)	c/a	B_0 (GPa)	$\Delta E^{bcc/hcp}$	$\Delta E^{fcc/hcp}$
	3.18	1.50	125	+0.326	+0.041
Zr	[3.15]	[1.62]	[102]	[+0.030]	[+0.027]
	(3.23)	(1.59)	(88)		

Table I: Calculated equilibrium properties compared to experiment (parenthesis) or FP-LMTO values [brackets]. a_0 is the equilibrium lattice parameter, c/a is the *hcp* axial ratio, and B_0 is the bulk modulus. $\Delta E^{bcc/hcp}$ and $\Delta E^{fcc/hcp}$ (in eV/atom) are energy differences between bcc and fcc phases and the equilibrium hcp phase.

	C_{11}	C_{33}	C_{44}	C_{66}	C_{12}	C^{rel}
Zr	219	183	28	74	71	34
	(158)	(174)	(37)	(47)	(63)	

Table II: Elastic constants in hcp phase for Zr (in GPa) compared to experimental harmonic values extrapolated at 0 K (parenthesis). $C_{11}^{rel}=C_{33}^{rel}=C_{66}^{rel}=-C^{rel}$ and $C_{12}^{rel}=+C^{rel}$.

The calculated elastic constants, shown in table II, satisfy the Born criterion, ensuring mechanical stability of the hcp phase. The deviation with respect to experiments is comparable to that of $ab-initio$ calculations for C_{33}, C_{44} and C_{12} (5%, 24%, and 12% respectively). The largest deviation, 57%, occurs in C_{66} where the calculated value is 74 GPa instead of 47 GPa. C_{11} displays an average result with a deviation of 38%. The RMS deviation over all calculated hcp elastic constants reaches 14 GPa, to be compared with 36 GPa found over all the hcp transition metals in reference [4]. The relaxation contribution takes the value of 34 GPa and leads to a ratio $C^{rel}/C_{66}^{(0)} = 31\%$ in agreement with the [10%-40%] range predicted by d-band models [10] for Zr.

CONCLUSIONS

We have presented a comparison between simple spd implementations of the $bond$ and $band$ formalisms of the two-center orthogonal tight-binding scheme. We have taken advantage of an $ab-initio$ data base of the band-structures to fit the hopping integrals and on-site energy differences. We have shown on the example of zirconium that a minimal set of parameters describing the exponential decay of the hopping parameters is capable of reproducing very well the band structures in three different structures. The fact that the energy difference between the hcp and the bcc phases changes by 0.54 eV when going from one model to the other outlines the importance of the energy references. Only the $bond$ model reproduces the stability of the hcp-phase. However the description of the on-site energy differences (fitted to constant values here) has to be improved to get a more quantitative agreement for the energy difference between the bcc and the hcp phases. The properties of the hcp-phase itself, including elastic constants, are in good agreement with experiments, making the present $bond$ model a good candidate for performing other calculations in hcp-Zr.

REFERENCES

1. F. Willaime and C. Massobrio, Phys. Rev. B **43**, 11653 (1991).
2. B. Legrand, Phil. Mag. A **52**, 83 (1985).
3. R. Ahuja, J.M. Wills, B. Johansson and O. Eriksson, Phys. Rev. B **48** 16269 (1993).
4. M. J. Mehl and D. A. Papaconstantopoulos, Phys. Rev. B **54**, 4519 (1996).
5. W. M. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989).
6. A.P. Sutton, M. W. Finnis, D. G. Pettifor and Y. Ohta, J. Pys. C **21**, 35 (1988).
7. M. Methfessel and A. T. Paxton, Phys. Rev. B **40**. 3616 (1989).
8. D. C. Wallace, *Thermodynamics of Crystals* (Wiley, New York, 1972).
9. A. T. Paxton, J. Phys. D **29**, 1689 (1996).
10. M. Nastar and F. Willaime, Phys. Rev. B **51**, 6896 (1995).

ENVIRONMENT-DEPENDENT TIGHT-BINDING MODEL FOR MOLYBDENUM

H. HAAS[1,2], C. Z. WANG[2], M. FÄHNLE[1], C. ELSÄSSER[3], K. M. HO[2]
[1] Max–Planck–Institut für Metallforschung, Heisenbergstraße 1, D-70569 Stuttgart, Germany
[2] Ames Laboratory and Department of Physics, Iowa State University, Ames, IA 50011, USA
[3] Max–Planck–Institut für Metallforschung, Seestraße 92, D-70174 Stuttgart, Germany

ABSTRACT

A transferable orthogonal tight–binding model for molybdenum is developed which goes beyond the traditional two–center approximation. The elements of the Hamiltonian matrix as well as the repulsive potential are allowed to depend on the environment. Several bulk, atomic defect and surface properties are calculated and compared with ab–initio data and experimental results to check the accuracy of the model.

INTRODUCTION

The original domain of tight–binding (TB) methods was the description of strongly covalent solids with special emphasis on Si and C. The application to metals is probably more demanding: in simple metals the s and p electrons are delocalized and the two–center approximation becomes questionable, and in transition metals there is a coexistence of delocalized s and p electrons and more localized d electrons.

The objective of this paper is to develop a TB parametrization for molybdenum which is as accurate and transferable as possible to see how far we can go with a TB description of this metal. Special emphasis will be given to the dependence of the matrix elements and the pair potential on the environment, very much in the sense of the method developed by Tang et al. [1] for C. The effective environment–dependence of orthogonal two–center Slater–Koster parameters may arise from (a) the explicit neglect of three–center matrix elements, (b) the use of an orthogonal basis set, which means that we are implicitly dealing with Löwdin orbitals whose Hamiltonian matrix elements depend on the structural environment and (c) the use of a minimal basis set. Concerning the pair potential $\phi(r_{\alpha,\alpha'})$, a contribution to the environment–dependence may arise from neglected many–body exchange–correlation terms in the double counting term of the Harris-Foulkes functional (which is the basis for the derivation of every TB model) as well as from that part of the pair–potential which implicitly accounts for the neglect of non–diagonal overlap matrix elements.

Another objective of the present paper is to develop this model in such a way that it is suitable for molecular-dynamics simulations which means that we will confine ourselves to interactions up to typically at most third-nearest neighbor interactions.

THE MODEL

We use a minimal basis set of one s, three p and five d atomic orbitals to construct the TB Hamiltonian. For an elementary metal we obtain 10 independent hopping parameters $V_{ll'm}$ and three intraatomic matrix elements (ϵ_s, ϵ_p and ϵ_d) whose environment–dependence is taken into account by $\epsilon_d = \epsilon_d^0 + \sum_{\alpha''} \Delta\epsilon_d(r_{\alpha\alpha''})$, $\epsilon_s = \epsilon_d + \epsilon_{s-d}^0 + \sum_{\alpha''} \Delta\epsilon_{s-d}(r_{\alpha\alpha''})$ and an analogous equation for ϵ_p. Here the quantities with superscript 0 denote the parts

independent of the environment and the $\Delta\epsilon_l$ are the environment–dependent contributions. In the following we adopt the same functional form for the distance-dependence of $V_{ll'm}$, $\Delta\epsilon_l$ and $\phi(r_{\alpha\alpha'})$, namely

$$f(r_{\alpha\alpha'}) = C_1 \exp(-C_2 r_{\alpha\alpha'})(1 - S_{\alpha\alpha'}) \ , \tag{1}$$

i.e., the primary distance-dependence is given by the exponential and the environment–dependence is modeled by the screening function introduced by Tang et al. [1],

$$S_{\alpha\alpha'} = \tanh 2\xi_{\alpha\alpha'} \ , \tag{2}$$

$$\xi_{\alpha\alpha'} = C_3 \sum_{\alpha''} \exp\left(-C_4 \frac{r_{\alpha\alpha''} + r_{\alpha'\alpha''}}{r_{\alpha\alpha'}}\right)^{C_5} \ . \tag{3}$$

There are 5 parameters C_i (with $C_2 \ldots C_5 > 0$) for the 14 functions $V_{ll'm}$ (10), $\Delta\epsilon_l$ (3) and ϕ (1), and the 3 parameters ϵ_d^0, ϵ_{s-d}^0 and ϵ_{p-d}^0. In order to reduce the number of parameters used for the fits we impose the universal ratios $V_{pd\sigma} : V_{pd\pi} = -\sqrt{3} : 1$ and $V_{dd\sigma} : V_{dd\pi} : V_{dd\delta} = (-6) : 4 : (-1)$ for the preexponential factors. Furthermore, we realized from the fits that $V_{pp\pi}$ is small and can be neglected without a noticeable change of the results. In contrast to a previous model [2], we now scale the distance between two atoms according to their coordination numbers using the scaling function introduced by Tang et al. [1], which adds two more parameters to the model (using this bond–length scaling considerably improves the TB total energies of the sc and A15 structure). Altogether the final model contains 55 parameters which are determined by fitting to the ab–initio band structures along lines of high symmetry in the Brillouin zone and the total energy for Mo in various crystal structures (sc,bcc,fcc) and for a variety of lattice parameters around the respective equilibrium lattice parameters, the experimental phonon frequencies at the points N, H, P of the Brillouin zone, the experimentally obtained elastic constant C_{44}, the vacancy formation energy obtained by the mixed–basis pseudopotential (MBPP) method and the unrelaxed (100) surface energy obtained by the MBPP method.

RESULTS

Fig. 1 shows the fitted total energy curves for sc, bcc and fcc Mo as well as the results for Mo in hcp and A15 structure, which were not included in the fit. Fig. 2 exhibits the fitted curves to the bandstructure in bcc Mo. Table 1 represents the TB results for bcc Mo for the equilibrium lattice constant a_0, the elastic constants C_{11}, C_{12} and C_{44}, the vacancy

	a_0 [a.u.]	C_{11} [Mbar]	C_{12} [Mbar]	C_{44} [Mbar]	E_V^f [eV]	E_I^f [eV]
TB	5.935	4.75±0.10	1.45 ± 0.10	0.99 ± 0.04	2.95	10.55
MBPP	5.926 [3]	—	—	—	2.90 ± 0.1 [3]	9.54 [4]
experiment	5.945 [5]	4.50 [6]	1.73 [6]	1.25 [6]	2.9 [7]	—

Table 1: Results of the TB model for the equilibrium lattice constant a_0, the elastic constants C_{11}, C_{12} and C_{44}, the vacancy formation energy E_V^f, and the formation energy E_I^f of an octahedral interstitial atom in a relaxed supercell containing 16 sites, in comparison with results from ab-initio MBPP calculations and experimental data.
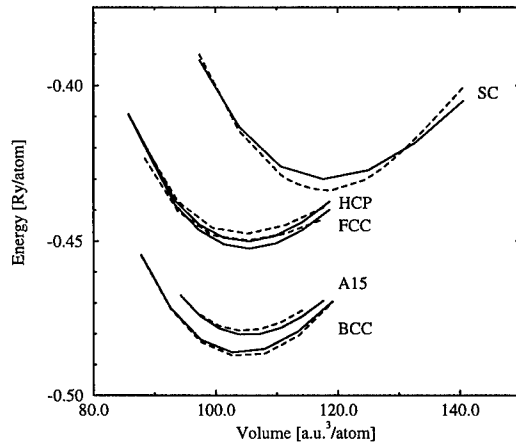
Figure 1: Total energy versus volume for sc, bcc, fcc and hcp Mo and for Mo in the A15 structure. The dashed and the full lines represent the TB and ab–initio data obtained by the linear-muffin-tin-orbital method in atomic-sphere approximation (LMTO-ASA), respectively.
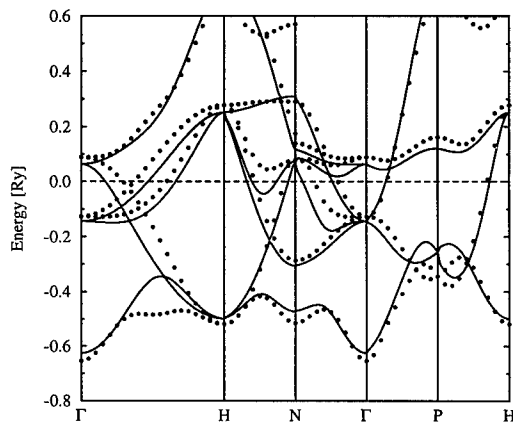
Figure 2: The TB band structure (dots) of bcc Mo at $a_0 = 5.8$ a.u. in comparison with the ab-initio LMTO-ASA band structure (solid lines).

329
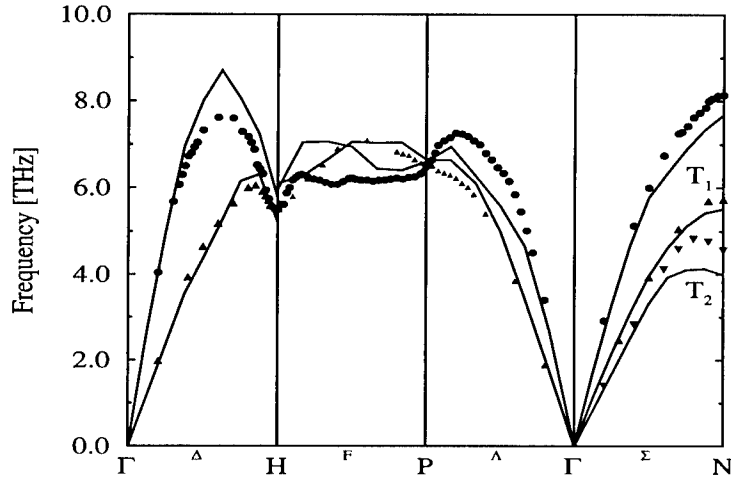
Figure 3: Comparison of the phonon frequencies in bcc Mo from the TB method (frozen–phonon calculation, full lines) and from inelastic neutron scattering [8] at $T = 296$ K (dots).

formation energy E_V^f for a relaxed supercell with 54 sites and the formation energy E_I^f for an octahedral interstitial atom in a relaxed supercell with 16 regular lattice sites, in comparison with experimental results and with results from the MBPP approach. The quantities a_0 and C_{44} were included in the fit, and therefore the comparison simply tests for the quality of the fit. In contrast, the results for C_{11}, C_{12}, for E_V^f in the relaxed supercell and for E_I^f are predictions of the TB model, which agree rather well with the data from experiments (C_{11}, C_{12}, E_V^f) and/or the MBPP calculation (E_V^f, E_I^f).

In Fig. 3 we compare the data from frozen-phonon type TB calculations with the experimental results [8] obtained by inelastic neutron scattering at 296 K. In spite of some quantitative discrepancies the qualitative agreement is satisfactory: the most important phonon anomalies in Mo, i.e, the low frequency of the H point phonon, the lowering of the T_2 mode when approaching the N point along the Σ line and the crossing of the longitudinal and the transversal phonon branches along the F line are correctly reproduced by the TB model.

To investigate the structure and electronic properties of the Mo (100) surface, we use a supercell containing seven layers. Fig. 4 and Fig. 5 display the bandstructure along high symmetry lines of the irreducible 2D Brillouin zone, calculated with the FLAPW method [9] and with the TB model, respectively. The agreement especially of the surface states close to the Fermi energy is promising.

Full relaxation of the $p(1 \times 1)$ slab yields a top–layer contraction of 7.9%, compared to 10.7% calculated with the MBPP method [10], and $(9.5 \pm 2)\%$ (Ref. [11]) respectively 11.5%
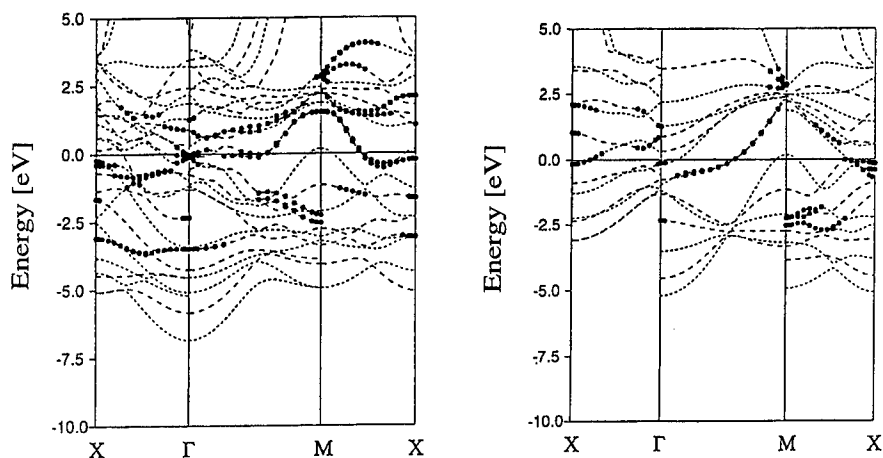
Figure 4: Energy bands for the Mo (100) surface for even (left) and odd (right) states with respect to the given symmetry lines, calculated with the FLAPW method [9]. Solid circles indicate surface states whose charge densities have more than 50% weight within the surface Mo layer.
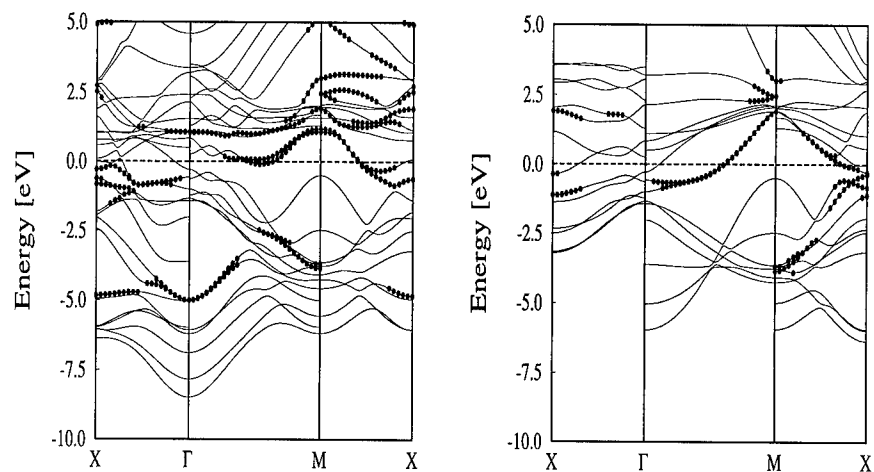
Figure 5: Energy bands for the Mo (100) surface for even (left) and odd (right) states with respect to the given symmetry lines, calculated with the TB model. Solid circles indicate surface states whose charges according to a Mulliken population analysis have more than 50% weight within the surface Mo layer.

(Ref. [12]) obtained by LEED experiments. For the surface energy of the relaxed $p(1 \times 1)$ surface our TB model yields a value of 1.89 eV per surface atom in satisfactory agreement with the MBPP value of 2.07 eV per surface atom [10]. Employing a $c(\sqrt{2} \times \sqrt{2})45°$ unit cell, we found that the (100) surface is stable with respect to displacements according to a M_1 surface phonon mode but unstable with respect to an M_5 surface phonon mode resulting in a $c(\sqrt{2} \times \sqrt{2})45°$ reconstruction of the (100) surface, which is in agreement with MBPP calculations [10].

SUMMARY

We developed a TB model for molybdenum which empirically incorporates the environment–dependence of interactions and demonstrated its transferability by comparing the TB results with ab–initio and experimental results for different atomic configurations.

Surface calculations with larger supercells are currently performed to find out whether the model is able to yield the experimentally observed $(7\sqrt{2} \times \sqrt{2})45°$ reconstruction [13]. In a future application we intend to use TB calculations to investigate the core structure of $(1/2) <111>$ screw dislocations in molybdenum.

ACKNOWLEDGMENTS

REFERENCES

1. M. S. Tang, C. Z. Wang, C. T. Chan and K. M. Ho, *Phys. Rev. B* **53**, 979 (1996).

2. H. Haas, C. Z. Wang, M. Fähnle, C. Elsässer and K. M. Ho, submitted to Phys. Rev. B.

3. B. Meyer and M. Fähnle, submitted to Phys. Rev. B.

4. B. Meyer, private communication.

5. Y. Waseda, K. Hirata and M. Ohtani, *High Temp. High Press.* **7**, 221 (1975).

6. F. H. Featherston and J. R. Neighbours, *Phys. Rev.* **130**, 1324 (1963).

7. R. Ziegler and H.-E. Schaefer, *Mat. Sci. Forum* **15-18**, 145 (1987).

8. B. M. Powell, P. Martel and A. D. B. Woods, *Can. J. Phys.* **55**, 1601 (1977).

9. S. C. Hong and J. I. Lee, in *International Conference on the Physics of Transition Metals* (World Scientific Publishing, Singapore, 1992), p. 524.

10. X. W. Wang, C. T. Chan, K. M. Ho and W. Weber, *Phys. Rev. Lett.* **60**, 2066 (1988).

11. L. J. Clarke, *Surf. Sci.* **91**, 131 (1980).

12. A. Ignatiev, F. Jona, H. D. Shih, D. W. Jepsen and P. M. Marcus, *Phys. Rev. B* **11**, 4787 (1975).

13. D. M. Smilgies, P. J. Eng and I. K. Robinson, *Phys. Rev. Lett.* **70**, 1291 (1993).

SEMI-EMPIRICAL TIGHT-BINDING PARAMETERS FOR TOTAL ENERGY CALCULATION IN ZINC

A. BERE, A. HAIRIE, G. NOUET and E. PAUMIER
Laboratoire d'Etudes et de Recherches sur les Matériaux, upresa-cnrs 6004
Institut des Sciences de la Matière et du Rayonnement
6 Bd du Maréchal Juin 14050 Caen Cedex France
Hairie@ismra.unicaen.fr

ABSTRACT

The semi-empirical tight-binding method is used to build up an interatomic potential in zinc. Using relaxed structures, the parameters are fitted to the lattice parameters, the elastic constants and the vacancy formation energy. The total energy calculation predicts the stability of the h.c.p. structure. The potential is used to calculate the energy of some extended defects : the basal stacking fault and two twin boundaries.

INTRODUCTION

The reliability of atomic scale simulation strongly depends on the interatomic potential. The simple empirical pair potentials are not able to describe all the characteristics of the interactions between the atoms. In many cases, it is necessary to take into account the electronic structure, using the density functional theory or the tight-binding approximation to build many-body potentials.

Numerous studies on various kinds of defects in zinc have been performed with pair potentials [1] but the more recent potentials are the phenomenological many-body ones proposed by Igarashi, Khantha and Vitek [2] and by Mikhin and de Diego [3]. They are fitted to several quantities, in particular to the non ideal value of the c/a ratio of the lattice parameters. They correctly predict the stability of the h.c.p. structure with respect to cubic ones.

Unfortunately, they both give a negative energy for the $(11\bar{2}2)$ twin boundary. Such a result predicting the stability of a defective crystal with respect to the perfect one is not satisfactory. It is our motivation to build a potential giving reasonable values of energy of extended defects in zinc. The high c/a ratio is thought to be a consequence of the electronic structure and this suggests to choose the semi-empirical tight-binding method.

After a short review of the two potentials cited above, the present paper explains how the semi-empirical tight-binding method is applied to zinc to give the energy of several extended defects as the basal stacking fault and the $(11\bar{2}2)$ and $(10\bar{1}2)$ twin boundaries.

MANY-BODY POTENTIALS

The potential of Igarashi et al. [2] is built in the Finnis-Sinclair [4] scheme for eight hexagonal metals. For each of them, the parameters are fitted to the two lattice parameters a and c, the cohesive energy, the unrelaxed vacancy formation energy and the elastic constants. The potentials successfully describe the mechanical and structural stability, the elastic constants and the phonon dispersion curves. However, in case of zinc, a very low value (0.4 mJm^{-2}) of the energy of the stacking fault in the basal plane is obtained. It is much lower than the estimates found in literature [5] (15 to 300 mJm^{-2}) and the result obtained by Legrand [6] in a pseudo-potential approach (35 mJm^{-2}).

Mikhin and de Diego [3] try to overcome this difficulty by using the second-moment of the tight-binding approximation to account for the d-electron interaction. Instead of the cubic splines used by Igarashi et al. [2], a screened Coulomb potential derived from pseudo-potential theory is used in the repulsive term and a combination of Gaussians in the attractive term. The parameters are fitted to the same quantities as for Igarashi et al. potential, and this leads to a correct description of all the quantities except C_{33}. The value of the stacking fault energy is found to be 10 mJm^{-2}, still rather low, even if it is improved with respect to Igarashi et al. value.

It may be noted that both potentials give further information about point defects properties as formation energy of self interstitials and vacancies. These two potentials seemed to be reliable enough to be applied to grain boundaries, but a problem arises with the $(11\bar{2}2)$ one. We found that its energy is slightly negative (-1.8 mJm^{-2}) with Igarashi et al. potential and an analogous result is known with Mikhin and de Diego potential [7]. Such a result is not satisfactory and seems contradictory to the mechanical and structural stability given by the two potentials [2,3]. Anyway, it leads us to propose another potential for computing the total energy of zinc based on the semi-empirical tight-binding (SETB) method.

APPLICATION OF SETB METHOD

The total energy of the atomic system under consideration is written as the sum of the electronic term U_{el} and the repulsive term U_R [8]:

$$U = U_{el} + U_R \qquad (1)$$

The electronic term describes the electron-ion interaction and is attractive. The second term U_R describes the short-range interaction between ions, and the cohesion of the material is due to the balance between the two terms.

Electronic energy

The electronic term U_{el} is obtained with the following method. The tight-binding hamiltonian operator of electrons is written in a basis formed by the atomic orbitals $\phi_\mu(r_i)$:

$$H = \sum_{\mu,i} \varepsilon_\mu \left|\phi_\mu(r_i)><\phi_\mu(r_i)\right| + \sum_{\mu,\nu} \sum_{i,j} V_{\mu\nu}(r_i r_j)\left|\phi_\mu(r_i)><\phi_\nu(r_j)\right| \qquad (2)$$

where $\phi_\mu(r_i)$ is the μ orbital of the atom i at position r_i and ε_μ is the atomic energy of this orbital independent on the atom position, $V_{\mu\nu}$ (r_i r_j) is the interaction matrix element between orbital μ of atom i and orbital ν of atom j, depending on length r and direction cosines of the vector joining atoms i and j. These matrix elements are expressed in terms of the hopping matrix elements in the nearest-neighbour tight-binding approximation of Harrison [9] for s and p orbitals, only. The matrix elements are multiplied by a cutoff function :

$$f_1(r) = \frac{1}{2} (1 - \tanh \frac{r - r_1}{\Delta r_1}) \qquad (3)$$

where r_1 is the cutoff radius and Δr_1 the length on which the cutoff takes place. The cutoff zone is established beyond the twelve nearest neighbours, by choosing $r_1 = 1.28a$ and $\Delta r_1 = 0.02$ a.

The electronic energy is given by :

$$U_{el} = 2 \sum_{i\alpha j\beta} H_{i\alpha j\beta} \, \rho_{j\beta i\alpha} \qquad (4)$$

where $H_{i\alpha j\beta}$ and $\rho_{i\alpha j\beta}$ are the elements of the hamiltonian and of the density matrix ρ between orbitals α of atom i and β of atom j. A O(N) method already developed [10, 11] is based on the method of Li, Nunes and Vanderbilt [12] to use the Lanczos Haydock recursion method. It allows a local calculation of the density matrix. With the cutoff radius r_1 given above and with a recursion level equal to twelve, the convergence of the method is satisfactory.

Repulsive energy

In the SETB scheme, the repulsive term in eq. (1) is usually written as a two body potential. Here, it is chosen as the product of two functions :

$$U_R = F(r) \, f_2(r) \qquad (5)$$

where $F(r)$ is a polynomial form in $r/r_a - 1$ (r_a being chosen as equal to the mean value of the distances between nearest - neighbours) given by :

$$F(r) = A_0 + A_1 \, (r/r_a - 1) + A_2 \, (r/r_a - 1)^2 + A_3 \, (r/r_a - 1)^3 \qquad (6)$$

and where $f_2(r)$ is a cutoff function of the same form as $f_1(r)$ but with possibly different characteristics :

$$f_2(r) = \frac{1}{2} \, (1 - \tanh \frac{r - r_2}{\Delta r_2}) \qquad (7)$$

r_2 is the cutoff radius and Δr_2 is the length on which the cutoff takes place. The effect of this cutoff is to limit the range of the repulsion between ion cores to take into account the screening by the electrons. The parameters r_2 and Δr_2 are chosen to give negligibility small values of $f_2(r)$ on the second neighbours. Each of the four parameters of $F(r)$ (Eq. 6) has a direct influence on a special property, which makes fitting procedure easier. The parameter A_0 is fitted to the vacancy formation energy, A_1 to the lattice parameters, A_2 to the bulk modulus and A_3 to the Grüneisen constants, which introduces some anharmonicity in the potential.

RELAXATION AND FITTING PROCEDURE

The equilibrium atomic configuration at zero temperature is that of minimum energy in which the sum of forces on each atom is zero. The calculation of the interatomic forces is the main reason of the high computation cost. Moreover, we found that our choice of basis and the use of the small localization region do not fulfill the conditions for applying Hellmann-Feynman theorem which leads to wrong results. Then, the interatomic forces are calculated by numerical gradients. Each atom is submitted to very small displacements (0.002 a), until the change in energy is less than 10^{-3} meV.

By using this relaxation procedure, the parameters of the repulsive term (Eq.6) are fitted to the perfect crystal properties. Then, the total energy of the crystal may be considered as a function of c/a ratio and by trial and error method, the c/a ratio giving the minimum energy is obtained.

Finally, with the values $r_a = a$, $r_2 = 1.28a$ and $\Delta r_2 = 0.02a$, for the lengths involved in Eq. 6 and Eq. 7 the best values of the other parameters in Eq. 6 are $A_0 = 2.35$ eV, $A_1 = - 8.47$ eV, $A_2 = 22.25$ eV and $A_3 = - 53.04$ eV. This best fit corresponds to a contraction of c axis leading to a c/a
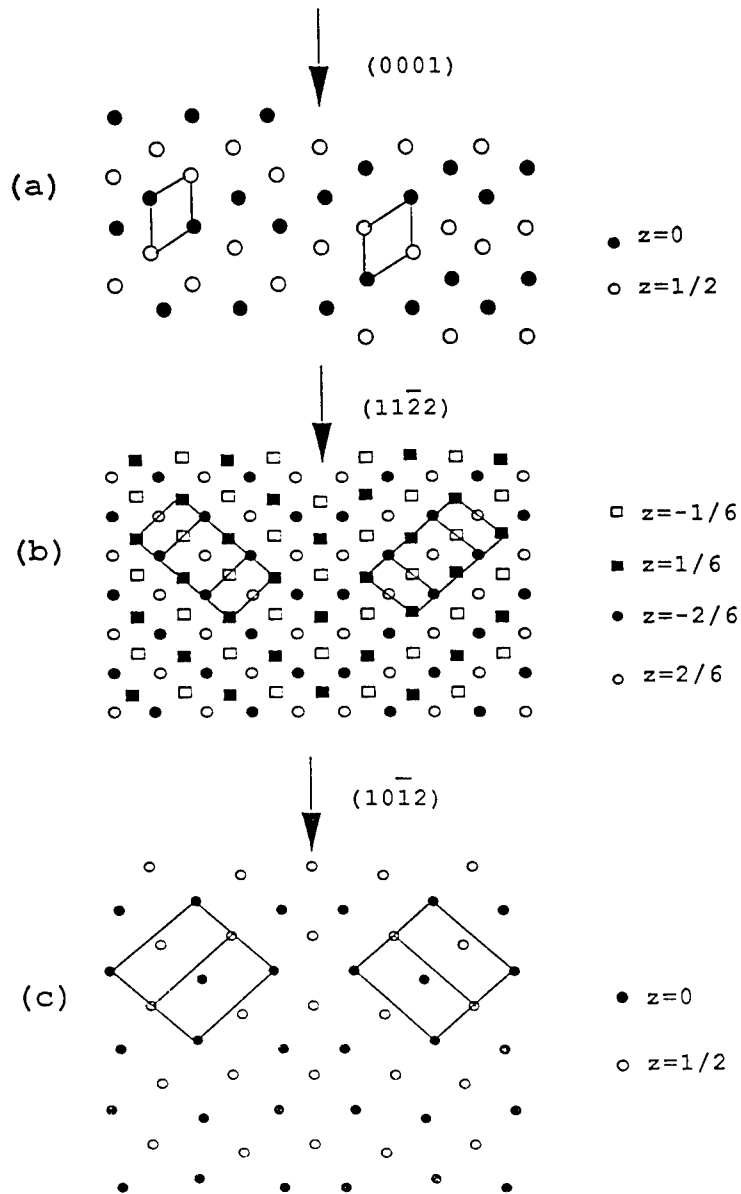
Fig. 1 : Projection of the relaxed atomic structure of the basal stacking fault along $[11\bar{2}0]$ (a), of the $(11\bar{2}2)$ twin boundary along $[1\bar{1}00]$ (b) and of the $(10\bar{1}2)$ twin boundary along $[\bar{1}2\bar{1}0]$ (c). Hexagonal unit cells are shown in outline, and the symbols distinguish atoms in different levels.

ratio equal to 1.801 (instead of 1.856). The vacancy formation energy is equal to 1.04 eV twice larger than the known experimental value (0.52 eV) [14].

RESULTS

The potential is now used to test the relative stability of different polytypes of zinc : three hexagonal structures (h.c.p. or 2H, 4H and 6H) and the f.c.c structure. They differ by their stacking periodicity : the h.c.p. structure is obtained by the stacking ABABAB and the f.c.c. one by ABCABC .... The 4H structure is obtained by the stacking ABCBABCB... characterized by a four layer periodicity. In the same way, the 6H structure is characterized by the stacking ABCACBABCACB ... with a six layer periodicity. Each structure is relaxed and its energy is calculated with the method above described. The results are given in table I with respect to the energy of the h.c.p. structure taken as the origin. These results show that the h.c.p. structure is the most stable one. Table I also gives the c/a ratio for each polytype.

Table I : Energy (in meV/atom) and c/a ratio of different polytypes.

Polytype	h.c.p.	4H	6H	f.c.c.
Energy	0	20.8	25.8	33.9
c/a	1.801	1.708	1.676	1.633

Another set of results is obtained for the relaxed structure of three planar extended defects : the basal stacking fault, the $(11\bar{2}2)$ and the $(10\bar{1}2)$ twins. The atomic positions in the stacking fault are shown in figure 1a, resulting from the stacking sequence ABABCACA .... As indicated above, its energy is not precisely known, and our result given in table II is between the theoretical one of Legrand [6] and the experimental estimates of Harris and Masters [5]. The atomic positions of the twins are given in figures 1b and 1c for $(11\bar{2}2)$ and $(10\bar{1}2)$, respectively. Both twins have been studied by Simon [15] with a pseudopotential and by Serra and Bacon [16] for all h.c.p. metals with six Lennard-Jones potentials. Moreover, the $(10\bar{1}2)$ has been studied theoretically by Hagège, Mori and Ishida [17] with a Lennard-Jones potential and by Braisaz, Ruterana, Lebouvier and Nouet [18] with Igarashi et al. potential and experimentally by Lay and Nouet [19] and by Braisaz, Ruterana, Nouet and Pond [20]. Table II gives the energy found in each case, either directly in the quoted papers or calculated from them. As usual it is impossible to compare the energy of the twin with experimental values, but our values seem statisfactory when compared to the other available ones. It may be noted that the energy of the $(10\bar{1}2)$ twin is larger

Table II : Energy of extended defects in mJm^{-2}, with respect to the perfect crystal (* calculated with Igarashi potential [2]).

Defect	Stacking fault	$(11\bar{2}2)$twin	$(10\bar{1}2)$ twin
Pseudopot [15]		290	1270
Pseudopot [6]	35		
Lennard Jones [17]			372.4
Many-body [2]	0.4 [2]	− 1.8*	125 [18]
Many-body [3]	10		
SETB	119.1	45.5	110

than the one of the $(11\bar{2}2)$. Serra and Bacon [16] have found the same order with two potentials but the opposite one (or nearly equal energies) with the other potentials.

Our method gives a better description of the electronic density of states than second moment method already used. This improvement leads to a better estimates of interface energies. Another way of improvement could be found in the pseudo-potential approach, but the computing cost could be higher than in the SETB method for the study of grain boundaries.

CONCLUSIONS

Using the semi-empirical tight-binding method, the energy of the relaxed structure of several extended defects in zinc is calculated. Contrary to the results obtained with recent phenomenological many-body potentials, the energy of $(11\bar{2}2)$ twin is positive. The results obtained for the $(10\bar{1}2)$ twin and for the basal stacking fault are satisfactory. The method should be improved by taking into account the d orbitals with the same method of relaxation.

REFERENCES

1. A. Serra, R. Pond and D.J. Bacon, Acta Metall. Mater. **39**, 1469 (1991) and literature quoted therein.
2. M. Igarashi, M. Khantha and V. Vitek, Phil. Mag. **B63** , 603 (1991).
3. A.G. Mikhin and N. de Diego, Phil. Mag. **A73**, 1211 (1996).
4. M.W. Finnis and J.E. Sinclair, Phil. Mag. **A50**, 45 (1984).
5. J.E. Harris and B.C. Masters, Proc. Roy. Soc. **A292**, 240 (1966).
6. B. Legrand, Phil. Mag. **B49**, 171 (1984).
7. A. Serra, private communication.
8. A.P. Sutton, M.W. Finnis, D.G. Pettifor and Y. Ohta, J. Phys. **C21**, 35 (1988).
9. W.A. Harrisson, Electronic Structure and the Properties of Solids (W.H. Freeman and Company, 1980).
10. A. Hairie, F. Hairie, B. Lebouvier and E. Paumier, Mat. Sci. Forum **207-209**, 105 (1996).
11. J. Chen, A. Béré, A. Hairie, G. Nouet and E. Paumier,Computational Materials Science **501** (1997).
12. X.P. Li, R.W. Nunes and D. Vanderbilt, Phys. Rev. **B47**, 10891 (1993).
13. R. Haydock, Solid State Physics (ed. F. Seitz and D. Turnbull, 1980) **35**, 215.
14. A. Seeger, Phil. Mag. **A64**, 735 (1991).
15. J.P. Simon, J. Phys. F : Metal Phys. **10**, 337 (1980).
16. A. Serra and D.J. Bacon, Phil. Mag. **A54**, 793 (1986).
17. S. Hagège, M. Mori and Y. Ishida, J. Phys. Cl **51**, 161 (1990).
18. T. Braisaz, P. Ruterana, B. Lebouvier and G. Nouet, Phys. Stat. Sol. **B191**, 267 (1995).
19. S. Lay and G. Nouet, Phil. Mag. **A70**, 261 (1994).
20. T. Braisaz, P. Ruterana, G. Nouet and R.C. Pond, Phil. Mag. **A75**, 1075 (1997).

THE ENVIRONMENT-DEPENDENT INTERATOMIC POTENTIAL
APPLIED TO SILICON DISORDERED STRUCTURES AND PHASE TRANSITIONS

MARTIN Z. BAZANT*, EFTHIMIOS KAXIRAS* and J. F. JUSTO**
*Department of Physics, Harvard University, Cambridge, MA 02138, bazant@cmt.harvard.edu
**Instituto de Fisica da Universidade de São Paulo, CP 66318, CEP 05315-970 São Paulo - SP, Brazil

ABSTRACT

The recently developed Environment-Dependent Interatomic Potential (EDIP) holds the promise of a new degree of transferability in describing bulk phases and defects of elemental covalent solids with a simple theoretically motivated functional form. Here we explore to what extent the environment-dependence of the model can extrapolate successes of the fitted version for Si for bulk defects to disordered phases, which involve local configurations very different from those used in fitting. We find that EDIP-Si provides an improved description of the metallic bonding bond angles of the liquid and is the first empirical potential to predict a quench directly from the liquid to the amorphous phase. The resulting amorphous structure is in closer agreement with *ab initio* and experimental results than with any artificial prepration method. We also show that melting of the bulk crystal and premelting of the (100)2×1 surface are reasonably well described by EDIP-Si in spite of its not being fit to any such properties.

INTRODUCTION

The availability of realistic and efficient interatomic potentials could have a major impact on the microscopic understanding of materials properties, for two basic reasons. The first is that simulations would be possible of complex processes requiring too many atoms or too long times to be feasible for simulation with a quantum-mechanical treatment of the electrons. The second benefit offered by a realistic interatomic potential is a conceptual framework in which to understand the complexities of chemical bonding. For example, the natural partitioning of total energy among the atoms and the distinction between radial and angular forces are very useful concepts for building intuition about atomic behavior in materials.

The difficulty in exploiting these benefits is that they are accompanied by serious drawbacks. The first is that there is no general and systematic method for deriving interatomic potentials that has demonstrated superiority over the ubiquitous "guess-and-fit" approach, which is to guess a functional form containing adjustable parameters and then to fit them to some set of experimental or *ab initio* data. Secondly, once one has an interatomic potential, its transferability is often in doubt. In other words, the potential cannot reliably describe structures different from those used in the fitting.

These problems have been most serious for covalently bonded solids, as exemplified by the prototypical case of silicon for which there are over 30 potentials in the literature, having a wide range of functional forms and fitting strategies but possessing comparable overall accuracy [1]. In spite of intense recent efforts, the simple, pioneering potentials of Stillinger-Weber (SW) and Tersoff continue to be the most popoular, albeit with well-documented limitations [1]. This experience suggests that the most successful approach is to build the essential physics into a simple functional form containing only a few adjustable parameters rather than brute-force fitting. The latter method, in our view also undermines the benefits of potentials described above: an overly complex functional form at the same time reduces the efficiency of force computation and hinders physical interpretation.

THE ENVIRONMENT-DEPENDENT INTERATOMIC POTENTIAL

In an effort to go beyond the practical limits of the guess-and-fit approach, we have provided *ab initio* theoretical guidance to motivate a functional form for interactions in covalent solids, which we call the Environment-Dependent Interatomic Potential (EDIP) [2]. Exact inversions of *ab initio* cohesive energy curves for silicon reveal global trends in bonding from covalent to metallic structures [3], and an analysis of elastic properties of the diamond and graphitic structures quantifies the mechanical properties of sp^2 and sp^3 hybrid covalent bonds [2]. We argue that the fundamental limitation of existing potentials is their inability to properly adapt bonding preferences to the local atomic environment, which is crucial for reproducing phase transitions.
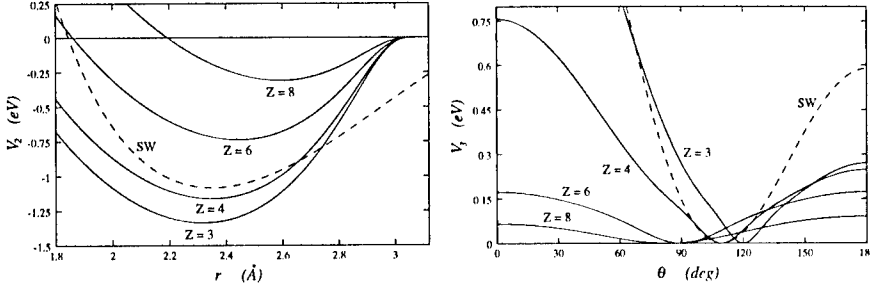
Figure 1: (Left) EDIP-Si two-body interaction $V_2(r, Z)$ as a function of separation r shown for coordinations 3, 4, 6 and 8. (Right) EDIP three-body interaction $V_3(r, r, \cos\theta, Z)$ for a pair of bonds of fixed length $r = 2.35$ Å subtending an angle θ, shown for coordinations 3, 4, 6 and 8. Each interaction is compared with the corresponding *environment-independent* term in the SW potential.

We now present the EDIP functional form and refer the reader to Ref. [2] for theoretical justification of all the terms. The energy of a configuration $\{\vec{R}_i\}$ is expressed as a sum over single-atom energies, $E = \sum_i E_i$, each containing radial and angular terms,

$$E_i = \sum_{j \neq i} V_2(R_{ij}, Z_i) + \sum_{j \neq i} \sum_{k \neq i, k > j} V_3(\vec{R}_{ij}, \vec{R}_{ik}, Z_i), \tag{1}$$

each dependent on the local atomic environment through an effective coordination number, $Z_i = \sum_{m \neq i} f(R_{im})$, where $f(r)$ is a cutoff function that determines the contribution to an atom's coordination from each of its neighbors,

$$f(r) = \begin{cases} 1 & \text{if } r < c \\ \exp\left(\frac{\alpha}{1 - x^{-3}}\right) & \text{if } c < r < a \\ 0 & \text{if } r > a \end{cases} \tag{2}$$

where $x = \frac{(r-c)}{(a-c)}$. The environment-dependent radial term takes the form of a (second moment) bond order model similar to embedded-atom potentials for metals,

$$V_2(r, Z) = A\left[\left(\frac{B}{r}\right)^\rho - e^{-\beta Z^2}\right] \exp\left(\frac{\sigma}{r - a}\right). \tag{3}$$

The strength of the attraction varies with coordination as dictated by inversion of cohesive energy curves [3]: as shown in Fig. 1, bond lengths get weaker and longer with increasing coordination, modeling sp^2 and sp^3 hybrid bonds and the covalent to metallic transition. The angular term essential for covalent bonding,

$$V_3(r, r', \theta, Z) = \lambda \exp\left(\frac{\gamma}{r - a}\right) \exp\left(\frac{\gamma}{r' - a}\right) H\left((\cos\theta + \tau(Z))Q_\circ^{1/2} e^{-\mu Z/2}\right), \tag{4}$$

$$H(x) = 1 - e^{-x^2} + \eta x^2, \tag{5}$$

also contains strong adaptation to the local environment through the the preferred angle $\cos^{-1}\tau(Z)$ and the variable strength of the angular function $e^{-\mu Z/2}$. The function $\tau(Z)$ is chosen (before fitting the potential) to interpolate between values suggested by theory ($\tau(4) = 1/3$, $\tau(3) = 1/2$, $\tau(6) = \tau(2) = 0$),

$$\tau(Z) = u_1 + u_2(u_3 e^{-u_4 Z} - e^{-2u_4 Z}), \tag{6}$$

with the parameters $u_1 = -0.165799$, $u_2 = 32.557$, $u_3 = 0.286198$, and $u_4 = 0.66$. As shown in Fig. 1, the angular function shifts its preferred angle to mimick covalent rehybridization for $Z \leq 4$ and softens considerably for overcoordinated, metallic environments.

Recently, we have presented a version of EDIP with 13 parameters fitted for bulk defects in silicon that displays remarkable transferability for bulk properties (elastic constants, bulk crystal structures, point
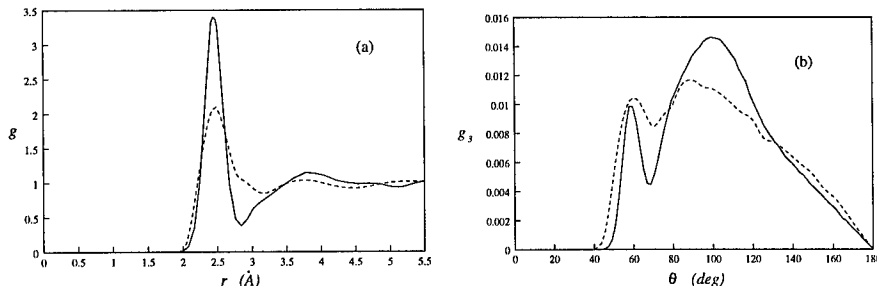
Figure 2: (Left) pair correlation function and (right) bond angle distribution for the liquid at $T = 1800$ K and $P = 0$ with EDIP-Si (solid lines) and an *ab initio* model [5] (dashed lines).

defects, concerted excahnge, stacking faults and dislocation core properties) [4]. The optimal parameter set we found is: $A = 7.9821730$ eV, $B = 1.5075463$, $\rho = 1.2085196$, $a = 3.1213820$ Å, $c = 2.5609104$ Å, $\sigma = 0.5774108$ Å, $\lambda = 1.4533108$ eV, $\gamma = 1.1247945$ Å, $\eta = 0.2523244$, $Q_o = 312.1341346$, $\mu = 0.6966326$, $\beta = 0.0070975$ and $\alpha = 3.1083847$. Since the fitted cutoff distance 3.1213820 Å is smaller than the SW cutoff, computing forces with our model is typically faster than with the SW potential (using an efficient algorithm [2]), making it possible to simulate thousands of atoms for millions of times steps in one day on typical workstation. Of course, at the moment this kind of efficiency is well beyond the reach of more accurate quantum-mechanical methods.

LIQUID PHASE

With the ability to simulate large systems for long times, it is natural to turn our attention to disordered structures and phases transitions. These situations contain a rich set of local bonding states from covalent (amorphous) to metallic (liquid) about which no information was included in our fitting procedure, thus constituting a stringent test of transferability for the model. Existing potentials have had considerable difficulty in simultaneously describing the crystalline, liquid and amorphous phases [1, 5, 6, 9]. In the case of the liquid, environment-independent potentials are not able to reproduce the bond angle distribution of the liquid[9], which deviates significantly from the tetrahedral angle.

We prepared a 1728-atom liquid sample with EDIP-Si potential at $T = 1800$ K and zero pressure in two ways, using standard techniques (Anderson piston, velocity rescaling, 5th order Gear predictor corrector integration). First, following a lattice start at 4000 K, the sample is cooled and equilibrated for about 50 ps. The same structure is also produced by heating a periodic bulk crystal, but the solid superheats before melting (an issue we address below). The latent heat of melting is 37.8 kJ/mol, in reasonable agrement with the experimental value of 50.7 kJ/mol, closer than the SW value of 31.4 kJ/mol [10]. The structural properties of the EDIP-Si liquid are shown in Fig. 2 and compared with the results of 64-atom *ab initio* simulations[5]. From the area under the first peak of $g(r)$, the coordination of the EDIP-Si liquid is 4.5, well below the experimental value of 6.5. Although this is unphysical, EDIP-Si offers a qualitative improvement in the bond angle distribution function $g_3(\theta, r_m)$. As shown in Fig. 2, EDIP-Si reproduces the auxiliary maximum at $\theta = 60°$, and is the first to do so, although the primary maximum is shifted toward the tetrahedral angle away from the *ab initio* most probable angle of $\theta = 90°$. It is important to emphasize that reasonable liquid properties are predicted by our model without any explicit fitting to the liquid phase. The reduced density and excess of covalent bonds may be artifacts of the short cutoff of our potential, which is appropriate for the covalently-bonded structures used in the fitting, but is perhaps too short to reproduce overcoordination in metallic phases like the liquid.

AMORPHOUS PHASE

Experimentally, amorphous silicon is known to form a random tetrahedral network, with long-range disorder and short-range order similar to that of the crystal [7, 8]. *Ab initio* simulations with 64 atoms
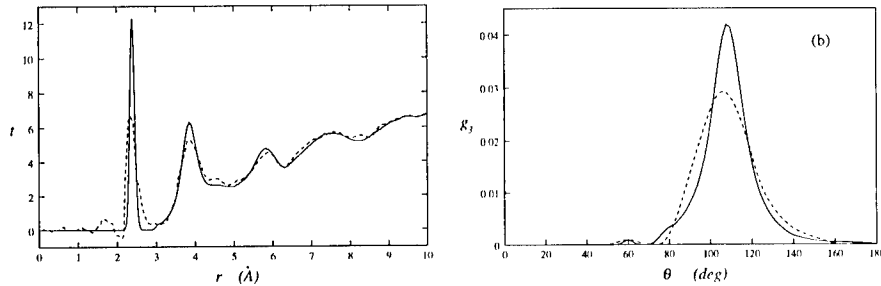
Figure 3: (Left) radial distribution function $t(r) = 4\pi\rho r g(r)$ for the amorphous phase at $T = 300$ K and $P = 0$ predicted by EDIP-Si (solid line) compared with the results of neutron scattering experiments[7]. (Right) bond angle distribution for EDIP-Si compared with *ab initio* results[6] (dashed lines).

	ρ_a	ΔH_{a-c}	$\bar{Z}$	$\bar{r}_1$	σ_{r_1}	$\bar{r}_2$	$\bar{r}_3$	$\bar{\theta}$	σ_θ
EDIP	0.04836	0.22	4.054	2.39	0.034	3.84	5.83	108.6	14.0
EXPT	0.044-0.054	< 0.19	3.90-3.97	2.34-2.36	0.07-0.11	3.84	5.86	108.6	9.4-11.0
LDA	-	0.28	4.03	2.38	0.079	3.84	-	108.3	15.5

Table I: Comparison of thermodynamic and structural properties of the present model (EDIP) for a-Si with (annealed) *ab initio*[6] and with (annealed) experimental[6, 7, 8] results. Shown are the density ρ_a in Å^{-3}, the excess enthalpy ΔH_{a-c} compared to the crystal in eV/atom, the coordination $\bar{Z}$, the mean $\bar{r}_1$ and deviation σ_{r_1} of the first neighbor distance in Å, the mean second $\bar{r}_2$ and third $\bar{r}_3$ neighbor distances in Å and the mean $\bar{\theta}$ and deviation σ_θ of the bond angles in degrees.

find almost 97% four-fold coordination [6]. An empirical potential would be invaluable in exploring larger system sizes and longer relaxation times than are feasible from first principles, but unfortunately no existing potential is capable of quenching directly from the liquid to the amorphous phase. Instead, empirical model liquids (like SW) typically transform into glassy phases upon cooling, characterized by frozen-in liquid structure [10]. Therefore, it has been impossible to simulate an experimentally relevant path to the amorphous structure (*e.g.* laser quenching [8]), and artificial preparation methods have been required to create large-scale amorphous structures [10, 11], but these do not involve realistic dynamics.

Remarkably, the EDIP-Si predicts a quench directly from the liquid into a high-quality amorphous structure. The phase transition is quite robust, since it occurs even with fast cooling rates. For example, quenching at -300 K/ps leads to a reasonable structure with 84% four-fold coordination. At much slower quench rates of -1 K/ps, an improved structure of 1728 atoms at $T = 300$ K and zero pressure is produced with almost 95% four-fold coordination. For comparison, the most common artificial prepration methods do not perform so well: the indirect SW method [10] predicts 81% four-fold coordination and the bond-switching algorimth of Wooten-Winer-Weaire [11] predicts 87%.

As shown in Fig. 3, the radial distribution function $t(r) = 4\pi\rho r g(r)$ is in excellent agreement with the results of neutron scattering experiments by Kugler *et. al.* [7] (using their experimental density $\rho = 0.054$ atoms/Å^3 for comparison). The persistence of intermediate-range order up to 10 Å captured by our model as in experiment is a strength of the empirical approach, since this distance is roughly the size of the periodic simulation box used in the *ab initio* studies [6]. Given the limited resolution of the experimental data, especially at small r (large q in the structure factor), the sharper first three peaks with our model may be interpreted as refinements of the experimental results. The bond angle distribution $g_3(\theta, r_m)$ shown in Fig. 3 (b) is narrowly peaked just below the tetrahedral angle, and also reproduces the small, well-separated peak at 60° observed in *ab initio* simulations [6] (unlike in previous empirical models). The peaks in both $g(r)$ and $g_3(\theta, r_m)$ are narrower and taller with EDIP-Si than with *ab initio* methods, which probably reflects the small system size and short times of the *ab initio* simulations compared to ours.
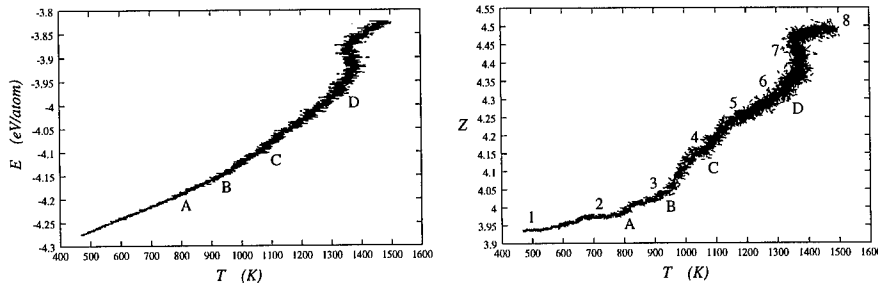
Figure 4: (Left) total energy and (right) average EDIP coordination number for a simulation of melting of a 3405-atom crystallite with (100)2×1 surfaces using EDIP-Si.
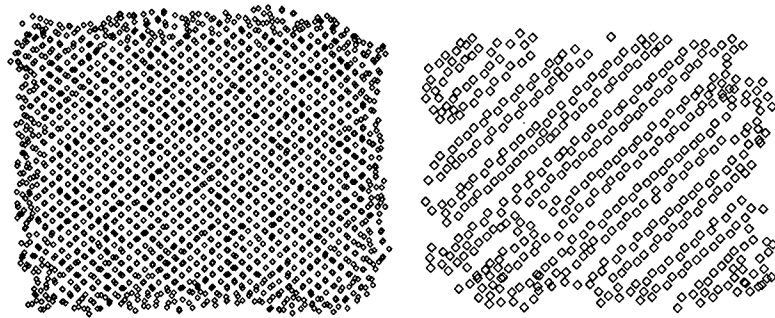
Figure 5: Two slices (12 Å thick) of the melting sample taken at stage 4, $T = 1000$ K. (Left) premelting of narrow (100) ledges. (Right) one of the two 7 nm × 7nm (100)2×1 surface just prior to the onset of premelting showing the presence of surface defects.

In Table I we summarize a detailed comparison of features of a-Si as obtained with our model and with ab-initio methods, against those revealed by experiment. Overall the agreement with experiment is very satisfactory, with the results of the present model somewhat closer to experimental values than ab-initio results as in the case of the enthalpy and the bond-length and bond-angle deviations. In summary, EDIP-Si faithfully reproduces the structure of amorphous silicon, following a realistic preparation procedure that starts with a liquid phase and cools it down without any artificial changes.

BULK MELTING AND SURFACE PREMELTING

Simulations of bulk melting with periodic boundary conditions have trouble in pinning down the melting point, as indicated by the spread of values reported for SW [1]. The problem is that in a periodic, defect-free crystal it is difficult for the system to nucleate the liquid phase. This difficulty is removed by introducing surfaces, which also allows us to study surface premelting. We performed simulations of 3405 crystal atoms arranged in a finite slab (7 nm × 7nm × 1.5 nm) terminated by (100)2×1 surfaces. The dimer reconstruction was artificially imposed on the original sample to avoid waiting for dimer rows to form spontaneously.

The sample was gradually melted by adding heat at a constant rate (by rescaling velocities) from 300 K to 1500 K in 2 ns (over 10 million time steps). Although a bulk periodic solid superheats and melts at around 2200 K, the finite slab undergoes a sharp, first-order transition at 1370 ± 20 K, as shown in Fig. 4. This is 20% below the experimental value of 1685 K, which is the closest agreement with the
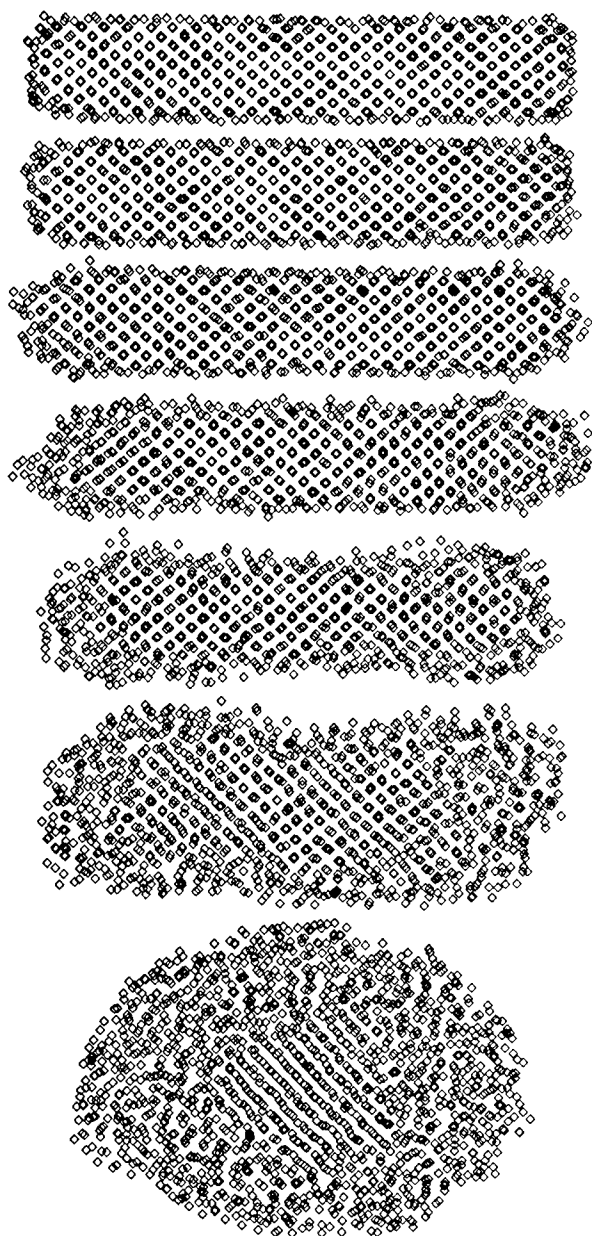
Figure 6: Snapshots taken at 0.2 ns (one million step) intervals of a simulation of a 3405-atom crystallite melting using EDIP-Si. The seven images from top to bottom correspond to the states 1–7 labeled in Fig. 4(b). The eighth state (not shown) is a spherical liquid drop.

melting point reported for any potential that was not explicitly fit to reproduce it[1]. The energy versus temperature curve shows some structure below the bulk melting point, but changes are much more clearly seen in the average EDIP coordination number, a convenient measure of local order. The system passes through 8 distinct states labeled in Fig. 4 and displayed in Fig. 6: (1) well-equilibrated (100)2×1 surfaces; (2) reconstructed edges exposing narrow (110) facets, give the sample a "beveled" look; (3) after passing through point A at 810 K, the edges and corners premelt, leaving faces intact; (4) after passing through point B at 950 K the narrow (2-4 dimers wide) (100)2×1 surfaces premelt, but the large surfaces do not. At this time, the large surfaces have some defects and mobile adatoms, as shown in Fig. 5; (5) after passing though point C at 1080 K, the large surfaces melt leaving two premelted monolayers encasing sample; (6) as the bulk melting point is approached, more layers premelt and surface and interface tension drive the sample to be more spherical by surface flow. The anisotropy of interfacial tension drives the interior crystallite to form (110) and (111) facets; (7) at $T = 1370$ K, added energy converts the bulk solid into liquid, leaving (8) a spherical liquid drop. Although the details of this simulation, primarily surface dynamics need to be more carefully validated, it demonstrates that EDIP-Si predicts premelting of the (100)2×1 surface about 200 K below the melting, consistent with experiment. Given our extensive tests for the liquid, amorphous and bulk defects, we may expect that the dynamics of the liquid-solid interface are fairly realistic by the usual standards.

CONCLUSION

In conclusion, EDIP-Si provides an improved description of disordered bulk structures and phase transitions. It is the first potential to predict the quench from liquid to amorphous and the structure is in excellent agreement with experimental and *ab initio* results. It also appears to describe several aspects of surface premelting, and predicts the bulk melting point to within 20%. These new results taken together with known successes for bulk defects suggest that EDIP-Si may be useful in applications such as solid and liquid phase epitaxial growth or radiation damage. Finally, we remark that the theoretical analysis which guided the development of EDIP-Si may be useful in extending the model to related covalent materials (Ge, C and possibly alloys).

ACKNOWLDGMENTS

REFERENCES

1. For a recent review, see H. Balamane, T. Halicioglu, and W. A. Tiller, Phys. Rev. B **46**, 2250 (1992).

2. M. Z. Bazant, E. Kaxiras and J. F. Justo, Phys. Rev. B **56**, 8542 (1997).

3. M. Z. Bazant and E. Kaxiras, Phys. Rev. Lett. **77**, 4370 (1996).

4. J. F. Justo, M. Z. Bazant, E. Kaxiras, V. V. Bulatov and S. Yip, submitted to Phys. Rev. B (1997).

5. I. Štich, R. Car and M. Parinello, Phys. Rev. B **44**, 4262 (1991); Phys. Rev. Lett. **63**, 2240 (1989).

6. I. Štich, R. Car, and M. Parrinello, Phys. Rev. B **44**, 11092 (1991).

7. S. Kugler, G. Molinar, G. Petö, E. Zsoldos, L. Rosta, A. Menelle and R. Bellissent, Phys. Rev. B **40**, 8030 (1989).

8. J. Fortner and J. S. Lannin, Phys. Rev. B **39**, 5527 (1989); Roorda, W. C. Sinke, J. M. Poate, D. C. Jacobson, S. Dierker, B. S. Dennis, D. J. Eaglesham, F. Spaepen and P. Fuoss, Phys. Rev. B **44**, 3702 (1991).

9. M. Ishimaru, K. Yoshida, and T. Motooka, Phys. Rev. B **53**, 7176 (1996).

10. J. Q. Broughton and X. P. Li, Phys. Rev. B **35**, 9120 (1987); W. D. Luedtke and U. Landman, Phys. Rev. B **37**, 4656 (1988).

11. F. Wooten, K. Winer and D. Weaire, Phys. Rev. Lett. **54**, 1392 (1985).

TRANSFERABLE TIGHT-BINDING APPROACH OF Si-H INTERACTIONS

Eunja Kim[a], Seung Mi Lee[b], and Young Hee Lee[b,c]
[a]Dept. of Physics, University of Nevada, Las Vegas, NV 89154
[b]Department of Semiconductor Science and Technology, Jeonju 561-756
[c]Dept. of Physics and Semiconductor Physics Research Center, Jeonju 561-756

ABSTRACT

We construct transferable tight-binding parameters of silicon-hydrogen interactions, reproducing the electronic energy levels and vibrational frequencies of the silane(SiH$_4$) molecule accurately. The potential function was rescaled with the exponential factor in order to ensure that the potential energy is smooth at an appropriate cut-off distance, which is very important in molecular-dynamics simulations. The parameters have been applied to other molecules and various surfaces such as hydrogenated Si(100) surfaces, for example, monohydride, dihydride, and (3x1) phase.

INTRODUCTION

The roles of hydrogen atoms in semiconductor materials are not only interesting physical phenomena but also very important to understand in device technology, because hydrogen is generally known to passivate surfaces and amorphous silicon(a-Si) by saturating the dangling bonds, changing the electrical and optical properties. [1] Especially, hydrogen plays an important role in layer-by-layer epitaxial growth of Ge on Si(100) surface. [2,3] In this case the surface morphology of Si overlayers strongly depends on surface H coverage. [4] Despite the fact of the abundant appearance in many H-Si systems, the physical mechanism underlying these phenomena is still far from being completely understood.

First principles approaches have been successfully applied to hydrogen in bulk Si and Si surfaces within the local density approximation(LDA). [5–7] Moreover, Car-Parrinello molecular-dynamics (CPMD) simulations [8] have also made a big step in further understanding of hydrogenated a-Si. [9] However, this method is limited to a small number of atoms, and the computational cost is still expensive for some dynamical properties. Although classical potentials have been proposed in hydrogenated a-Si [10], due to the low computational cost, the reliability cannot be guaranteed since it does not include the electronic information in total energy calculation.

Therefore, several empirical tight-binding (TB) models have been proposed. [11,12] Since the TB scheme considers only the valence electrons in the electronic structure calculations and the overlap integrals are empirically fitted to the first principles calculations or experiments, both the accuracy and efficiency have been achieved. Although the TB model by Min et el. [12] takes account of the anharmonic effect by implementing the universal binding energy curve in the fitting procedure, the simple inverse scaling law of the distance still causes

347

undesirable effects particularly in describing hydrogen dynamics and absorption-desorption processes. Therefore, a new model is generally required to have a smooth potential function behavior at an appropriate cutoff distance in a typical MD simulation.

Goodwin, Pettifor, and Skinner (GPS) [13] have proposed a transferable TB total energy scheme by introducing an exponential factor in the distance scaling of the silicon system. This has been successfully applied to many complex systems. [15] Now, in this paper, we propose a new TB model for H-Si system by introducing an exponential factor in the distance scaling similar to the GPS scheme such that the potential cutoff is smoothened so as to apply to more complex systems. We expect this model to be transferable to many complex systems, such as hydrogen dynamics in hydrogenated a-Si and the absorption-desorption reaction in chemical vapor deposition process or at hydrogenated Si surfaces.

FITTING PROCEDURE FOR H-Si

The total energy in an empirical TB scheme can be written as

$$E_{tot} = \sum_{n,k} \varepsilon_{n,k} + U. \tag{1}$$

Here the first term represents the band structure energy which can be constructed from the TB Hamiltonian H_{TB} and the second term is the repulsive ion-ion interactions. The diagonal terms of H_{TB} are equal to the corresponding atomic energy levels and the off-diagonal elements are constructed via the Slater and Koster empirical TB scheme [16].

For the Si-Si interaction, the traditional GPS total energy scheme has been introduced. [13,14] In this scheme, the elements of H_{TB} matrix are determined by

$$V_\alpha^{Si-Si}(r/r_0) = V_\alpha^{Si-Si}(1)(r_0/r)^n exp\{n(-(\frac{r}{r_c})^{n_c} + (\frac{r_0}{r_c})^{n_c})\} \tag{2}$$

where the V_α^{Si-Si} (1) are the hopping integrals at equilibrium distance r_0. The repulsive interaction term also has a similar form:

$$U^{Si-Si}(r/r_0) = U^{Si-Si}(1)(r_0/r)^m exp\{m(-(\frac{r}{r_c})^{m_c} + (\frac{r_0}{r_c})^{m_c})\}. \tag{3}$$

The total energy from this model well reproduces the universal binding energy curves of various phases generated by the first principles calculations [17]. The main idea of this scheme is to improve the transferability by introducing an exponential factor in the distance scaling, in contrast with the previous inverse square law by Chadi [18].

For the H-Si interaction, we adopt a scheme similar to the GSP. The respective parameters are shown in Table I. In this system, the energy levels and the vibrational frequencies of silane (SiH$_4$) were used to fit those parameter. The significance of this formulation is that the potential change is smooth at an appropriate cutoff distance, as required for realistic MD simulations, and the quantitative accuracy is improved.

TABLE I. The TB potential parameters of H-Si system.

$V_{ss\sigma}$(1) (eV)	$V_{sp\sigma}$(1) (eV)	r_0 (Å)	r_c (Å)	n	n_c	m	m_c	U (1) (eV)
-3.5535	5.088	1.48	2.186	1.9877	13.269	2.255	13.269	3.01

RESULTS AND DISCUSSION

Figure 1 shows the total energy curve of the H-Si interaction. One can see that both the band structure energy and the repulsive potential energy smoothly approach zero at the cutoff distance (2.186 Å) and the total energy shows the equilibrium at 1.48 Å with the binding energy 3.73 eV, as compared to the experimental value 3.355 eV.
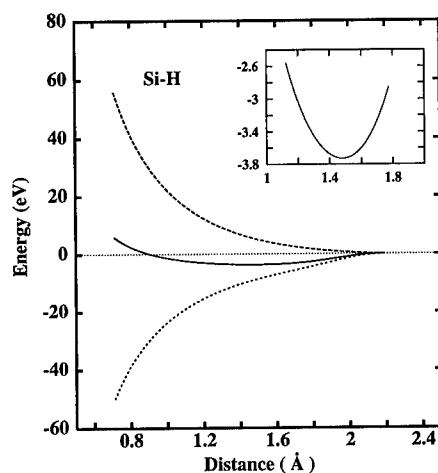
FIG. 1. The total energy curves of H-Si interaction. The band structure energy, repulsive potential energy (dashed lines) and the total energy (solid line) are smooth at the cutoff distance 2.186 Å.

The energy levels and the vibrational frequencies of the SiH_4 molecule have been calculated using this model as shown in Table II. In order to calculate the vibrational spectra, the velocity-velocity autocorrelation function is calculated, and then the Fourier transformations are also taken. For this purpose we run MD with a microcanonical ensemble at room temperature for 25 picoseconds, achieving a resolution of 0.001 meV.

TABLE II. The energy levels and the vibrational frequencies of SiH_4 molecule. The values in parentheses are from the experiment.(Ref.[12])

	Eigenvalues			Frequencies		
levels	present TB	Min et al.[a]		modes(meV)	present TBMD	Min et al.[a]
a_1^+ (eV)	-18.08(-18.23)	-18.23		symmetric bending	121(121)	122
t_2^+ (eV)	-12.60(-12.7)	-12.71		asymmetric bending	108(113)	103
t_2^- (eV)	-0.53(-5.4)	-1.93		symmetric stretching	276(270)	276
a_1^- (eV)	-3.34(-4.1)	-2.84		asymmetric stretching	284(271)	

[a] Reference [12]. [b] Reference [12] and references therein.

The energy levels of the antibonding states are not well described with a minimal basis set in the TB scheme. However, in the H-Si system, the antibonding states are far away from the Fermi level and do not strongly affect the bonding characteristics of the system. Although we sacrificed our antibonding states a little bit compared to Min et $al.$'s model, we improved the asymmetric bending frequency from 103 to 108 meV where the bending frequencies are solely determined from the change of $V_{sp\sigma}$ (H-Si). The difference of $V_{ss\sigma}$ and $V_{sp\sigma}$ determines the accuracy of the asymmetric bending mode in this case. (Our $V_{sp\sigma}$ (H-Si) value is a little larger than Min et $al.$'s value.)

We now apply this tight-binding model of the Si-H interaction to different molecules such as Si_2H_6 and Si_2H_4 and confirm the transferabilty of this model. [19] For example, the Si-Si bond in the fully relaxed Si_2H_6 molecule is 2.31 Å, not 2.35 Å, in this model. This can be understood as the contraction of the Si-Si bond from 2.35 to 2.31 Åbecause the electronegativity of the Si atom is a little larger than that of the H atom. The H-Si bond length becomes longer from 1.48 to 1.49 ÅThe bond angle between H atoms is 107 ° and the bond angle between Si and H atom is 111.8 °, as compared to the tetrahedral angle 109.5 °. These values are in good agreement with electron diffraction experimental results [21]. The calculated energy levels are -10.49, -12.15, -12.79, -16.74, and -18.50, which are also in better agreement with experiments than Min et $al.$'s results. [19] Moreover, the bending modes (113 meV, 105 meV) are in excellent agreement with experiment (112 meV, 105 meV) which are important in identifying the various hydrogen bonding species in many real systems.

Recently, Lee et $al.$ [22] has applied this model to the hydrogenated Si (100) surface. In this study 72 Si atoms, with 24 H atoms to saturate the dangling bonds of the bottom Si atoms, have been used to reproduce the hydrogenated (3x1) phase. Among the 6 Si layers, two layers of Si atoms from the bottom and the bottom H atom layer have been fixed to mimic bulk Si. Figure 2 shows fully relaxed geometries in terms of the hydrogen coverage. H-terminated Si(100) surfaces can be classified by H coverage as monohydride, dihydride, and the (3x1) phase.

When one monolayer(ML) of hydrogen atoms covers on dimerized Si(001) surface, this surface is called a monohydride. Although the starting Si surface is the asymmetric dimerized surface, as the exposed dangling bonds are completely saturated by H atoms the charges of the dimer atoms are compensated by hydrogen atoms resulting in a symmetric dimerization, as shown in Fig. 2(a). The dimer bond length is reduced from 2.45 Å to 2.39 Å after hydrogen chemisorption, close to the bulk bond length of 2.35 Å in good agreement with the pseudopotential result of 2.40 ÅThe Si-H bond length becomes 1.50 Å shorter than the pseudopotential result of 1.54 Å [23]. It seems that hydrogen atoms not only saturate the dangling bonds but also release strains near the surface, in good agreement with previous observation [24].

As the hydrogen coverage increases from 1 ML to 2 ML, hydrogen atoms break into the dimer bond and a (2×1) phase is changed to a (1×1) phase, the so called dihydride surface, as shown in Fig.2(b). The bond angle between two hydrogen atoms with the same Si atom is 102.4 °, similar to 102 ° found by pseudopotential calculation [23]. This angle is smaller than the tetrahedral bond angle of 109.5 ° due to the H-H repulsive interactions of adjacent hydrogen atoms of the different Si atoms. (For the H-H interactions a new tight-binding model has been introduced. [19]) The Si-H bond length of dihydride is slightly shorter than that of monohydride, showing a similar trend to the pseudopotential results.

For dihydride, several possible models have been suggested, including the canted model. [23,25,26] In order to test the stability of each suggested structure, the configurations are redistributed to the approximate position, and then fully relaxed by the SD method. The canted model, which has a small rotation of the H_2 groups, is found to be an energetically favorable structure, although the total energy difference from an ideal dihydride is negligible. Our TB scheme predicts the similar structure to the pseudopotential results qualitatively but not quantitatively due to the H-H interactions, fitted to H_2 molecule. [19] In general the H-H interaction on Si surface might be very different than that of H-H in the H_2 molecule, since each H atom makes a bond with different surface Si atoms. Unfortunately, it is beyond our scheme at this moment.
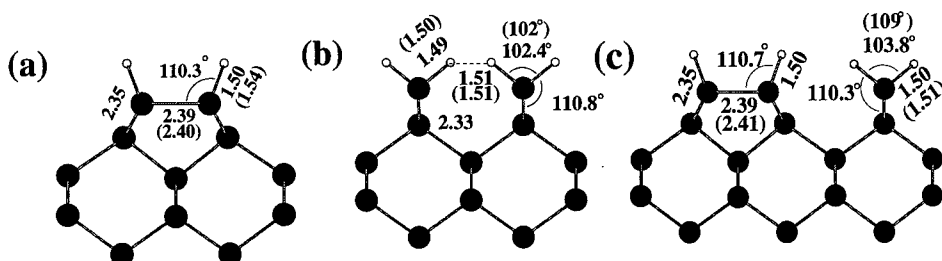
FIG. 2. (a) Fully relaxed monohydride upon one monolayer(ML) coverage, (b) ideal dihydride upon two ML coverage, and (c) fully relaxed (3×1) phase upon 1.3 ML coverage. The values in parentheses are those of pseudopotential calculations.[19] The bond lengths are in units of Å The dotted line in (b) simply indicates the distance between hydrogen atoms of adjacent Si sites.

When hydrogen coverage is in intermediate ML as 1.3 ML, the monohydride phase appears alternately with the dihydride phase along the $[1\bar{1}0]$ direction, the so called (3×1) phase as shown in Fig.2(c). Each monohydride and dihydride domain which consists of the (3×1) phase has similar behavior to its individual ideal structure shown in Fig. 2(a) and (b) respectively. The obtained bond lengths and bond angles are in good agreement with the pseudopotential calculation in parentheses [23]. The θ_{H-H} is slightly increased due to the farther distance, compared with that of an ideal dihydride, between the adjacent H atoms of different Si atoms.

CONCLUSION

A transferable TB scheme for the H-Si system has been constructed. The hopping integrals are empirically fitted to the SiH_4 energy levels and vibrational frequencies for Si-H interactions, respectively. The application of this method to other molecules such as Si_2H_6 and Si_2H_4 shows good tranferability for our model. Hydrogenated Si(100) surface such as monohydride, dihydride, and (3x1) phase, were excellently described by this Si-H interaction. We expect that this model can be applied to more realistic MD simulations

such as the absorption-desorption mechanism and diffusion mechanism of H atoms in Si(100) surfaces.

ACKNOWLEDGMENTS

REFERENCES

[1] For instance, see *The Physics of Hydrogenated Amorphous Silicon II*, edited by J. D. Joannopoulos and G. Lucovsky (New York:Springer Verlag 1984).
[2] A. Sakai and T. Tatsumi, Appl. Phys. Lett. **64**, 52(1994).
[3] R. Tsu, H. Z. Xiao, Y. W. Kim, M. A. Hasan, H. K. Birndaun, J. E. Green, D. S. Lin, and T. C. Chiang, J. Appl. Phys. **75**, 240(1994).
[4] M. Copel, R. M. Tromp, Phys. Rev. Lett. **72**, 1236 (1994).
[5] K. J. Chang and D. J. ChadiPhys. Rev. Lett. **60**, 1422(1989).
[6] C. G. Van der Walle, Y. Bar-Yam and S. T. Pantilides, Phys. Rev. Lett. **60**, 2761(1988).
[7] S. Jeong and A. Oshiyama, will be appeared at Phys. Rev. Lett.(December 1, 1997)
[8] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471(1985).
[9] F. Buda, G. L. Chiarotti, R. Car, and M. Parrinello, Phys. Rev. Lett. **63**, 294 (1989).
[10] N. Mousseau and J. Lewis, Phys. Rev. B**41**, 3702(1990).
[11] D. C. Allan and E. J. Mele, Phys. Rev. Lett. **53**, 826(1984).
[12] B. J. Min, Y. H. Lee, C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev. B **45**, 6839(1992).
[13] L. Goodwin, A. J. Skinner, and D. G. Pettifor, Europhys. Lett. **9**, 1701(1989).
[14] C. Z. Wang, C. T. Chan, and K. M. Ho, Phys. Rev.B **45**, 12227 (1992).
[15] E. Kim and Y. H. Lee, Phys. Rev. B**51**, 5429 (1995).
[16] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498(1954).
[17] M. T. Yin and M. L. Cohen, Phys. Rev. B**26**, 5668(1982).
[18] D. J. Chadi, Phys. Rev. Lett. **41**, 1062(1978); Phys. Rev. B**29** 785(1984); Phys. Rev. Lett. **59**, 1691(1988).
[19] E. Kim, Y. H. Lee, and J. M. Lee, J. Phys.: Condens. Matter **6**, 9561(1994).
[20] B. Beagley, A. R. Conrad, J. M. Freeman, J. J. Monaghan, B. G. Norton, and G. C. Holywell. J. Mol. Struct. **11**, 371(1972).
[21] D. Horowitz and W. A. Goddard, J. Mol. Struct. **163**, 207 (1988).
[22] S. M. Lee and Y. H. Lee, Surf. Sci. **347**, 329 (1996).
[23] J. E. Northrup, Phys. Rev. B **47**, 10032(1993).
[24] Z. H. Lu, K. Griffiths, P. R. Norton, and T. K. Sham, Phys. Rev. Lett. **68**, 1343(1992).
[25] S. Ciraci, R. Butz, E. M. Oelling, and H. Wagner, Phys. Rev. B **30**, 711(1984).
[26] X. M. Zheng and P. V. Smith, Surf. Sci. **279**, 127(1992).

NEGATIVE CAUCHY PRESSURE WITHIN THE
TIGHT-BINDING APPROXIMATION

D. NGUYEN-MANH *, D.G. PETTIFOR *, S. ZNAM **, V. VITEK **
*Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, U.K.
** Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, Pennsylvania, 19104-6272, U.S.A.

ABSTRACT

It is well-known that the Embedded Atom Method (EAM) predicts positive Cauchy pressures for cubic metals if physically-motivated embedding functions are used. Suprisingly, even if the angular character of the covalent bonding is included within an orthogonal or non-orthorgonal Tight-Binding (TB) description, the Cauchy pressure for most elemental and binary metallic systems remains positive. We describe the results of a detailed breakdown of the different contributions to the Cauchy pressure within the Harris-Foulkes approximation (HFA) to density functional theory. We show that negative values of the Cauchy pressure for both elemental transition metals such as Ir and binary intermetallics such as Ti_3Al, $TiAl$ and $TiAl_3$ are well reproduced by the HFA. We argue that the negative Cauchy pressure (NCP) arises namely from the environment dependence of the local TB orbitals which leads to both environment-dependent bonding integrals and overlap repulsion. We discuss a particular functional form for overlap repulsion which leads to NCP and compare it with different fitting schemes proposed recently in TB theory.

INTRODUCTION

Development of intermetallic compounds such as the titanium-aluminides as advanced engineering materials for use at high temperature in aircraft engines has attracted considerable attention in recent years [1]. These intermetallics have many attractive properties but a challenging problem remains their low-temperature brittleness. A possible origin of their brittleness could reside in the peculiar property of their elastic constants, namely, a *negative Cauchy pressure* (i.e. C_{12}-C_{44} < 0 for cubic materials, or C_{13}-C_{44} < 0 and C_{12}-C_{66} < 0 for hexagonal ones) [2]. This reflects important features of the nature of the bonding at the atomistic level. If the bonding can be described by pair-wise potentials, then the Cauchy pressure would be zero. If the bonding is more metallic in that spherical atoms are embedded in the electron gas of the surrounding neighbors, the Cauchy pressure would be positive as described by the EAM potentials [3]. A NCP seemed to imply that angular dependent many-body forces are playing a crucial role in intermetallic compounds. As the tight-binding (TB) model is the simplest scheme within a quantum mechanical framework for describing the bonding of materials, there was much hope that the NCP could be modelled successfully by a new many-body order N potential for the bond order, the Bond-Order Potential (BOP)[4]. Unfortunately, applications of these BOPs for intermetallic compounds have found that it is usually impossible to fit BOPs with physically motivated TB parameters that achieve NCP (see for example, [5]). The purpose of this paper is to elucidate the physical origins of this complex problem and to find a way to improve the BOPs for predicting of materials with NCP's properties.

RESULTS FROM THE TIGHT-BINDING BOND (TBB) MODEL

Within the two-center orthogonal TBB model, the binding energy can be written in

the form [4]

$$U = U_{rep}^{pp} + U_{bond} \tag{1}$$

where U_{rep}^{pp} is a semi-empirical repulsive *pairwise* contribution

$$U_{rep}^{pp} = \frac{1}{2} \sum_{i,j \neq i} \Phi(|\mathbf{R_i} - \mathbf{R_j}|) \tag{2}$$

and the attractive bond energy is given by

$$U_{bond} = \sum_{il} \int^{E_F} (E - E_{il}) n_{il}(E) dE = U_{band} - \sum_{il} E_{il} N_{il}. \tag{3}$$

U_{band} is the band-structure energy, E_{il} is (2l+1)-fold degenerate atomic energy level on atom i, n_{il} and N_{il} are the l-projected local density of states (DOS) and number of electrons, respectively. When evaluating the elastic moduli, we apply a small homogeneous strain tensor, $\epsilon_{\alpha\beta}$. At equilibrium, the equilibrium conditions are

$$\sigma_{\alpha\beta} = \sigma_{\alpha\beta}^{pp} + \sigma_{\alpha\beta}^{b} = 0 \tag{4}$$

where $\sigma_{\alpha\beta}$ is the average stress in the system, and $\sigma_{\alpha\beta}^{pp}$, $\sigma_{\alpha\beta}^{b}$ are respectively their pair-wise and bond components. The second order elastic moduli

$$C_{\alpha\beta\gamma\delta} = C_{\alpha\beta\gamma\delta}^{pp} + C_{\alpha\beta\gamma\delta}^{b} \tag{5}$$

should be all positive definite. Following equations (4) and (5), the Cauchy pressures using Voigt's notation are

$$C_{12} - C_{44} = C_{12}^{b} - C_{44}^{b} + \frac{1}{2}\sigma_{11}^{b} \tag{6}$$

for *cubic symmetry* and

$$C_{13} - C_{44} = C_{13}^{b} - C_{44}^{b} + \frac{1}{4}(\sigma_{33}^{b} + \sigma_{11}^{b})$$

$$C_{12} - C_{66} = C_{12}^{b} - C_{66}^{b} + \frac{1}{2}\sigma_{11}^{b} \tag{7}$$

for *tetragonal symmetry*. Equations (6) and (7) show that within the TBB model the Cauchy pressures are determined not only by the bond energy elastic constant contributions but also by its corresponding *stress tensor components* via the equilibrium condition (4). We have applied equation (6) for cubic fcc-Ir and (7) for tetragonal $L1_0$-TiAl where the orthogonal two-centre TB parameters are deduced from the ab-initio TB-LMTO method which gives a good features of the electronic DOS [6]. Our calculations have been made for different TB basis sets; in particular we use the down-folding method in TB-LMTO to deduce the d-band parameters for the transition metal (Ir) and the d-p parameters for the transition metal aluminide (TiAl) [7]. Our results for the Cauchy pressure within the TBB approximation are presented in the Table 1. We see that the total Cauchy pressure within the TBB model remains *positive* in contradiction with experiment for Ir and TiAl.

It is very clear from our calculations that the stress contributions in equations (6) and (7) are mainly responsible for the predicted positive values. The origin of this effect is physically understandable because of the decreasing of the (negative) bond energy as the

Table 1: Cauchy pressure (in GPa) within the TBB model calculated with different orbital basis sets.

Ir-fcc	(d)	(spd)	TiAl-L1$_0$	(d-p)	(spd)	TiAl-L1$_0$	(d-p)	(spd)
$C_{12}^b - C_{44}^b$	-10	-32	$C_{13}^b - C_{44}^b$	21	-6	$C_{12}^b - C_{66}^b$	20	28
$(1/2)(\sigma_{11}^b)$	38	336	$(1/4)(\sigma_{11}^b + \sigma_{33}^b)$	51	283	$(1/2)(\sigma_{11}^b)$	52	290
$C_{12} - C_{44}$	28	304	$C_{13} - C_{44}$	72	277	$C_{12} - C_{66}$	72	318
Exp.	-13	-13	Exp.	-27	-27	Exp.	-5	-5

distance is reduced. The larger the TB basis set (for example, on going from d to spd for Ir), the larger bond stress contribution results as can be seen in Table I. We made more tests by including the effect of the crystal-field contribution in order to improve the TB on-site energy description. In this case, the results for fcc-Ir(spd) did show that the stress contribution is reduced but the total Cauchy pressure still remains positive (177 GPa instead of 304 GPa from Table 1). Inclusion of non-orthogonality does not appear to improve the situation. A recent non-orthogonal TB fit to Si still predicts a strong positive CP of 31 GPa compared to the NCP of -16 GPa [8]. We conclude that even though the TBB model can give an accurate account of the band structure, it is *unable to provide the correct sign of the Cauchy pressure* within equation (1).

RESULTS FROM THE HARRIS-FOULKES APPROXIMATION (HFA)

To gain insight into the origin of the NCP, we need to understand the behavior of the elastic constants within the HFA from which the TBB model has been derived [9]. The HFA can be obtained from the variational principle of Kohn-Sham density functional theory (DFT)[10]. Taking the input charge density as the superposition of *free atomic charge densities* comprising the core density n_c and valence density n_v, the total binding energy can be written as [10]:

$$U^{HFA}[n^{in}] = \sum_{i,occ} E_{iv} - \frac{1}{2} \int \int d\mathbf{r} d\mathbf{r}' \frac{n_v(\mathbf{r})n_v(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{1}{2} \sum_{i,j \neq i} \frac{Z_{vi}Z_{vj}}{|\mathbf{R_i} - \mathbf{R_j}|} - E_v^{atom}$$

$$- \int d^3r \, n_v(\mathbf{r})\mu_{xc}(\mathbf{r}) + U_{xc}[n_v] + \delta T_c \tag{8}$$

where the first term is the band energy and the next three terms are equal to a change in the electrostatic energy, comprising the double-counting electrostatic energy, the core-core interaction energy, and the negative energy of the valence electrons in the free atomic state. In the second line of equation (8), the first two terms represent the change in the exchange and correlation energy whereas the last term δT_c is a correction to the total kinetic energy due to perturbed atomic cores. It is well-known that the TBB model approximates all terms in (8) except the bond energy by a sum of pair potentials. Our HFA calculations were carried out using the Full-Potential Liniear Muffin-Tin Method (FP-LMTO) [11] in which the input atomic charge density was multiplied by a Fermi-factor cutoff suggested by Finnis [12]. Our HFA results of Cauchy pressure for cubic fcc-Ir and fcc-Rh, tetragonal DO$_{22}$-TiAl$_3$ and L1$_0$-TiAl and hexagonal DO$_{19}$-Ti$_3$Al are shown in Table 2.

It is gratifying to see that the HFA can predict good values of the elastic constants. Moreover, the *negative* Cauchy pressures of all those materials are resonable well reproduced, except for the CP of C_{12}-C_{66} for TiAl and TiAl$_3$. Our further HFA calculations

Table 2: Elastic constant (in GPa) within the Harris-Foulkes appriximation

-	Ir	Ir	Rh	Rh	$TiAl_3$	$TiAl_3$	TiAl	TiAl	Ti_3Al	Ti_3Al
-	HFA	EXP	HFA	EXP	HFA	EXP	HFA	EXP	HFA	EXP
C_{11}	633	590	469	422	218	217	195	186	197	183
C_{33}	-	-	-	-	216	217	207	176	220	225
C_{44}	261	262	221	194	94	92	113	101	68	64
C_{66}	-	-	-	-	143	116	74	77	67	47
C_{12}	228	249	205	192	101	57	79	72	62	89
C_{13}	-	-	-	-	44	45	72	74	65	62
$C_{12} - C_{44}$	-32	-13	-16	-2	-	-	-	-	-	-
$C_{13} - C_{44}$	-	-	-	-	-49	-47	-41	-27	-3	-2
$C_{12} - C_{66}$	-	-	-	-	-42	-59	5	-5	-5	22

for bcc-Mo, diamond-Si, B1-TiC and C11b-MoSi$_2$ show that the Harris-Foulkes functional does indeed provide the correct sign of the Cauchy pressure.

From a detailed examination of the different energy terms in equation (8) we see that the the band energy and the electrostatic contribution including the double-counting part give *very large cancelling components* to the NCP (see Table 3). This large cancellation between the band energy and the electrostatic energy was one of the main reasons for developing the TBB model in which the large shifts in the on-site energies implicit in the *band* energy were removed and grouped with the negative electrostatic contributions to give much smaller contributions from the *bond* energy and electrostatic energy [13].

A detailed analytic study of the hydrogen molecule within the TBB description of the binding energy led automatically to a *very small* net electrostatic contribution at the equilibrium nearest neighbour distance within the HFA (see Fig.2 of [14]). The failure of the TBB expression, eq. (1) to yield NCPs resides in the fact that the orbitals used are *strictly tightly bound*, i.e. interactions extend out to only nearest neighbour atoms. This requires that the orbitals are not *unscreened* atomic orbitals but are orbitals that have been *screened* by the local atomic environment, either through the application of chemical pseudopentential theory (see [14] and references therein) or use of screened structure constants within TB-LMTO theory [6]. We consider how this affects the Cauchy pressure in the next section.

INCLUSION OF THE ENVIRONMENT DEPENDENCE

Table 3: Different energy contributions to the Cauchy pressure (in GPa) for TiAl from the Harris-Foulkes functional (eq. (8)).

U^{HFA}	C_{11}	C_{33}	C_{44}	C_{66}	C_{12}	C_{13}	$C_{13} - C_{44}$	$C_{12} - C_{66}$
Band	71	74	-397	-471	8	61	439	479
Electrostatic	134	144	497	549	87	25	-473	-462
Exchange-correlation	5	7	4	2	-5	-3	-6	-7
Core kinetic correction	-14	18	-9	-5	-9	-11	-2	-5
Total HFA	195	207	113	75	79	72	-41	5
Exp.	186	176	101	77	72	74	-27	-5

The environment dependence effects both the repulsive and bonding contributions in eq. (1) [14]. However, for the case of elemental transition metals, the valence d orbitals are not strongly screened by the local environment because they are fairly localised (see Fig. 4 of [15b]). Therefore, in this paper we consider the environmental dependence of the overlap repulsion only, since it arises in large part from the valence sp electrons in transition metals that we squeezed into the core region under strong influence of unsaturated covalent d bonds (see Fig.4b of [16]). In principle, the short-range part of repulsive interaction can be approximated by the Yukawa form

$$\Phi(\lambda_{ij}R_{ij}; \lambda_{ij}R_c) = \frac{A}{R_{ij}}exp(-\lambda_{ij}(R_{ij} - 2R_c)) \tag{9}$$

where R_c is the core radius. For simplicity, we shall neglect the explicit core dependence and write the repulsive energy as a sum over environmentally-dependent potentials of the form

$$U_{rep}^{env} = \frac{1}{2}\sum_{i,j\neq i} \Phi(\lambda_{ij}R_{ij}) \tag{10}$$

with the following definitions

$$\lambda_{ij} = \frac{1}{2}(\lambda_i + \lambda_j), \lambda_i = \lambda_\infty + \delta\lambda_i$$
$$\delta\lambda_i = f_i = f(\rho_i) = [\sum_{k\neq i}\rho(R_{ik})]^{1/\gamma} \tag{11}$$

where $\rho(R_{ik})$ is a pair-wise function. Calculating the general form of the elastic constants from $U=U_{rep}^{env}+U_{bond}$ is a 'wet towel job', but finally we can get a relatively simple analytic expression for the Cauchy pressure. For example, in the case of cubic symmetry, we obtain

$$C_{12} - C_{44} = C_{12}^b - C_{44}^b + \frac{1}{2}\sigma_{11}^b +$$
$$+\frac{1}{2V}\sum_i(\sum_{j\neq i}P_{ij}\frac{X_{ij}^2}{R_{ij}})(\sum_{k\neq i}\frac{Y_{ki}^2}{R_{ki}}\rho'(R_{ki})) + \frac{1}{2V}\sum_j(\sum_{i\neq j}P_{ji}\frac{X_{ij}^2}{R_{ij}})(\sum_{k\neq j}\frac{Y_{kj}^2}{R_{kj}}\rho'(R_{kj})) \tag{12}$$

where X_{ij} and Y_{ij} are the cartesian projection of R_{ij},

$$P_{ij} = f_i'(\Phi_{ij}' + \lambda_{ij}R_{ij}\Phi_{ij}'') + [f_i''\sum_{k\neq i}R_{ik}\Phi_{ik}' + \frac{1}{2}(f_i')^2\sum_{k\neq i}R_{ik}^2\Phi_{ik}'' + \frac{1}{2}f_i'f_j'R_{ij}^2\Phi_{ij}'']\rho'(R_{ij}). \tag{13}$$

$\rho'(R_{ij})$, f_i', f_i'', Φ_{ij}', Φ_{ij}'' denote the full derivatives of the corresponding functions with respect to their argument. Comparing with equation (6) the second line terms in equation (12) represent an additional contribution coming from the environment dependence of the new repulsive potential (10). For the case of fcc-Ir, after fitting the d-band model with experimental values of lattice constant and bulk modulus we find that C_{11}=569 GPa, C_{12}=272 GPa and C_{44}=284 GPa. The result gives C_{12}-C_{44}=-12 GPa which is in excellent agreement with the experimental value for the NCP. Importantly, we also find that in order to get a NCP the value of γ in equation (11) must be $\gamma > 1$.

Finally, it is not difficult to show the relation between our proposed functional form for the overlap repulsion (eqs.(10)-(11)) with other recently proposed environmentally-dependent TB schemes. For instance, if we suppose $\Phi(\lambda_{ij}R_{ij}) = Aexp(-\lambda_{ij}R_{ij})$ and $\rho(R_{ik}) = Cexp(-\mu R_{ik})$ then it is easy to justify that as in [15]

$$F(R_{ij}) = C_1exp(-C_2R_{ij})(1 - S_{ij}) \tag{14}$$

provided the screening function takes the functional form $S_{ij} \approx \delta\lambda_{ij}R_{ij}$.

Also we can write $U_{rep}^{env}=(U_{rep}^{env}-U_{rep}^{pp})+ U_{rep}^{pp}=U_{rep}^{ac}+U_{rep}^{pp}$, where U_{rep}^{ac} is a correction due to environment dependence to the pair-wise potential. Then in relation with the effective medium theory (EMT) [17], we have $U_{rep}^{ac}=\sum_i[\sum_{j\neq i} \Phi^{pp}(R_{ij})]^{\frac{1}{\gamma}}$, where $\gamma=1+\eta_1/\eta$ provided we take $AR_{ij}\exp(-\lambda_\infty R_{ij})$ as a constant. Note that within the EMT, $\eta > 0$ and $\eta_1 > 0$.

ACKNOWLEDGMENTS

REFERENCES

1. See Intermetallic compounds, vol. 1 & 2, eds. J.H. Westbrook and R.L. Fleischer, John Willey & Son, (1996).

2. A.H. Cottrel, in Proc. European Conf. Advanced Materials, p. 2, Cambridge, (1991).

3. M.S. Daw and M.T. Baskes, Phys. Rev. B., **29**, 6443, (1984).

4. D.G. Pettifor, Phys. Rev. Lett., **63**, 2480, (1989).

5. K.E. Tan, A.M. Bratkovsky, R.M. Harris, A.P. Horsfield, D. Nguyen-Manh, A.P. Sutton and D.G. Pettifor, Modelling Simul. Mater. Sci. Eng., **5**, 199, (1997).

6. O.K Andersen, O. Jepsen and M. Sob, in Electronic Band Structure and its Application, ed. by M. Yussouff, p.1, Springer-Verlag, Heidelberg-New York, (1987).

7. D. Nguyen-Manh, A. Bratkovsky and D.G. Pettifor, Phil. Trans. Royal Society of London, **351**, 529, (1995). D. Nguyen-Manh, unpublished (1997).

8. N. Bernstein and E. Kaxiras, to be published (1997).

9. A.P. Sutton, M.W. Finnis, D.G. Pettifor and Y. Ohta, J. Phys. C, **21**, 35, (1988).

10. J. Harris, Phys. Rev. B, **31**, 1770, (1985); W.M.C. Foulkes and R. Haydock, Phys. Rev. B, **39**, 12520, (1989).

11. M. Methfessel, Phys. Rev. B, **38**, 1537,(1988)

12. M. Finnis, J. Phys.: Condens. Matter (1991)

13. D.G. Pettifor, Commun. Phys. **1**, 141, (1976).

14. A.J. Skinner and D.G. Pettifor, J. Phys.: Condens. Matter, **3**, 2029, (1991).

15. M.S. Tang, C.Z. Wang, C.T. Chan and K.M. Ho, Phys. Rev. B, **53**, 979, (1996); H. Hass, C.Z. Wang, M. Fahnle, C. Elsasser and K.M. Ho, to be published (1997).

16. D.G. Pettifor, J. Phys. F, **8**, 219, (1977).

17. K.W. Jacobsen, Comments Cond. Mat. Phys., **14**, 129, (1988); M. Sob and V. Vitek, in Stability of Materials, eds. by A. Gonis, P.E.A. Turchi and J. Kudrnovsky, p. 449, New York, Plenum Press, (1996).

CHARACTERIZATION OF INTERATOMIC POTENTIALS BY
A CALCULATION OF DEFECT ENERGY

Y. KOGURE, M. DOYAMA
Teikyo University of Science & Technology, Uenohara, Yamanashi 409-01, Japan

ABSTRACT

Potential functions used in molecular dynamics simulations for metals are characterized through a calculation of the third-order elastic constants, the Gruneisen parameters, and the molecular dynamics simulation of point defects. The Lennard-Jones potential and the embedded atom method potentials for noble metals (Cu, Ag, Au) are characterized by using a common program code.

INTRODUCTION

The empirical interatomic potential is widely used in the molecular dynamics simulation for the structure of crystal and amorphous, defects, phase transition, and so on. In metals the interaction between atoms has non-central character due to the valence electrons, which causes a deviation from the Cauchy relation of elastic constants. Recently, the embedded atom method potential (EAM) has been developed by many authors to express the many body interaction in metals [1-5]. These potential functions are compared on the points of the reproducibility of the physical properties and the adaptability to the molecular dynamics simulations.

POTENTIAL FUNCTIONS

Lennard-Jones Potential

The Lennard-Jones (L-J) potential is a typical two body potential based on the van der Waals force and is able to reproduce the properties of rare gas solid. The potential has a simple functional form and has been widely applied to the fundamental simulation studies.

The potential is described by two parameters, ϵ and σ, and is often adopted in the simulation of metals because it has a simple functional form. The interaction energy for two atoms separated by r is expressed as

$$E(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right]. \tag{1}$$

In the molecular dynamics simulation the potential function is truncated at a finite distance, r_c, and the parameters depend on r_c. The values of r and r_c are normalized by the nearest

neighbor distance r_0 through this paper. A truncation, $r_c = 1.9$, is chosen in the present study , in which an atom in a fcc crystal interacts with 42 atoms. The potential parameters ϵ and σ are determined by fitting the potential to the experimental values of the cohesive energy, the lattice parameter, and the second-order elastic constants on the basis of the least mean square method. Hereafter the potential function is referred as 'L-J'.

General form of EAM potential

The EAM potentials are commonly expressed as

$$E_i = F(\rho_i) + \frac{1}{2}\sum_{j \neq i} \phi(r_{ij}), \tag{2}$$

where, ρ is the electron density function and it is a sum of the density of the neighbor atoms labeled by j.

$$\rho_i = \sum_{j \neq i} f(r_{ij}) \tag{3}$$

A Variety of forms for $F(\rho)$, $f(r)$ and $\phi(r)$ have been proposed.

A-T-V-F (Ackland et al.) potential [3]

Finnis and Sinclair have developed a N-body potential for transition metals based on the tight binding theory, which is equivalent to the EAM potential by Baskes et al. The method was applied to fcc metals by Ackland et al. The embedded function $F(\rho)$ is given by

$$F(\rho) = -\sqrt{\rho}, \tag{4}$$

and the functions $\phi(r)$ and $f(r)$ are

$$\phi(r) = \sum_{k=1}^{6} a_k (r_k - r)^3 H(r_k - r), \tag{5}$$

$$f(r) = \sum_{k=1}^{6} A_k (R_k - r)^3 H(R_k - r). \tag{6}$$

Here, $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x > 1$, namely, each term in the summation is truncated at r_k's or R_k's. The maximum values of the truncation distance is $1.2347a_0$, where a_0 is the lattice parameter.

O-J (Oh and Johnson) potential [4]

Oh and Johnson have developed a potential based on the exponential function.

$$F(\rho) = a(\rho/\rho_e)^n - b(\rho/\rho_e), \tag{7}$$

$$\phi(r) = \phi_{old}(r) - \phi_{old}(r_c) - g(r)\phi'_{old}(r_c)/g'(r), \tag{8}$$

$$f(r) = f_{old}(r) - f_{old}(r_c) - g(r)f'_{old}(r_c)/g'(r), \tag{9}$$

$$f_{old}(r) = f_e \exp[-\beta r - 1], \quad \phi_{old}(r) = \phi_e \exp[-\gamma r - 1], \quad g(r) = 1 - \exp[\delta(r - r_c)], \tag{10}$$

where, $a, b, n, \beta, \gamma, \delta$ and ϕ_e are the fitting parameters. The truncation distance is chosen to be $r_c = 1.9$. The potential is referred as 'O-J'.

E(r) [eV]
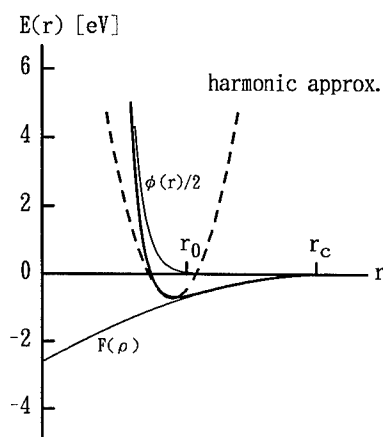
Fig.1 EAM ptential energy for interacting two atoms. Two thin lines show contributions from $F(\rho)$ and $\phi(r)$, separately. The dashed line demonstrates harmonic approximation, schematically.

D-K (Doyama and Kogure) potential [5]

Recently, the authors have developed a new potential functions.

$$F(\rho) = D\rho \ln \rho, \quad \rho = \sum_{j \neq i} f(r_{ij}). \tag{11}$$

where, $F(\rho)$ is the embedding energy for i-th atom, ρ is the electron density, and r_{ij} is the distance between i-th and j-th atom. The functions $\phi(r)$ and $f(r)$ are

$$\phi(r) = A(r_c - r)^2 \exp(-c_1 r), \tag{12}$$

$$f(r) = B(r_c - r)^2 \exp(-c_2 r), \tag{13}$$

where, r_c is a cut off distance of the potential. These functions contain five parameters A, B, C_1, C_2, and D. They are determined by fitting the potential to experimental values of physical properties for crystals.

ANHARMONICITY

Third-order elastic constants

An example of the EAM potential energy between two atoms is shown as a function of separation distance in Fig. 1. The position and the energy of the potential minimum and are related to the lattice parameters and the cohesive energy of the crystal. The second derivative of the potential function is related with the second order elastic constants, which are the coefficients of linear stress-strain relation due to the harmonic approximation. The anharmonicity is a deviation from the linear stress-strain relation and can be expressed by the third-order elastic constants (TOE). The constants are calculated from the third derivative of the energy E about the strain η. There are six elements of TOE in fcc crystals.

Table I   Calculated elastic constants. (in units of 100 GPa)

	L-J	A-T-V-F	O-J	D-K	Exp.
c_{11}	1.492	1.756	1.734	1.781	1.684
c_{12}	0.936	1.176	1.202	1.206	1.214
c_{44}	0.936	0.716	0.738	0.714	0.754
C_{111}	-21.545	-59.900	-9.882	-13.322	-12.71
C_{112}	-11.804	-21.899	-4.708	-5.740	-8.14
C_{123}	0.165	-8.574	-0.010	-0.491	-0.50
C_{144}	0.165	-9.223	-0.407	-0.391	-0.03
C_{166}	-11.804	-21.998	-4.426	-5.589	-7.80
C_{456}	0.165	-8.961	0.059	-0.032	-0.95

$$C_{ijklmn} = \frac{1}{\Omega}\frac{\partial E}{\partial \eta_{ij}\eta_{kl}\eta_{lm}}, \tag{14}$$

where, Ω is the atomic volume. The calculated TOE's for copper are summarized and compared with the experimental values (Exp.) in Table I. It can be seen that the potential functions expressed by a power function r^n's (L-J and A-T-V-F) give larger values of TOE than the experiments.

Grüneisen parameter

Although the third-order elastic constants express the elastic anharmonicity completely, available experimental data are rather limited. On the contrary, the Grüneisen parameter expresses the anharmonicity by a single value, and it can be obtained from the elastic and thermal measurement. The Grüneisen parameter for most of elements and many compounds are available in literature [6].

According to the higher order elasticity theory, the mode Grüneisen parameter can be calculated from the second- and the third-order elastic constants.

$$\gamma_{ij}(\text{elastic}) = -(1/w)[c_{abik} + c_{aubk}U_uU_i + c_{aubi}U_uU_k$$

$$+C_{aubvik}U_uU_v]N_aN_b, \tag{15}$$

$$w = c_{aubv}U_uU_vN_aN_b,$$

where, N_i is the propagation direction of the elastic wave, and U_i is the polarization vector. By taking average for sound propagation direction and the polarization (longitudinal and transverse wave), the Grüneisen parameters of noble metals (Cu, Ag, Au) have been calculated and the results are summarized in Table II.

The elastic Grüneisen parameter is also determined experimentally from the nonlinear volume-pressure relation,

$$\Delta V/V = aP + bP^2, \tag{16}$$

$$\gamma(\text{elastic}) = \frac{b}{a^2} - \frac{2}{3}. \tag{17}$$

Table II Calculated and experimental Grüneisen parameters.

	L-J	A-T-V-F	O-J	D-K	Exp.(elastic)	Exp.(thermal)
Cu	4.92	11.76	1.37	2.01	2.00	1.97
Ag	4.46	12.50	1.44	1.37	2.29	2.46
Au	5.47	12.62	1.57	1.38	2.22	3.09

The thermal Grüneisen parameter is derived from a relation

$$\gamma(\text{thermal}) = \frac{\alpha V}{B \cdot C_v}, \tag{18}$$

where, α is the thermal expansion coefficient, V is the atomic volume , B is the compressibility, and C_v is the heat capacity at constant volume. The experimental values of the Grüneisen parameter, E(elastic) and E(thermal), are also shown in Table II. The calculated Grüneisen parameters by the potential L-J and A-T-V-F are too large compared with the experimental values, whereas the values are even smaller in the other two potentials. The origin of the large difference in the anharmonicity may be attributed to the functional forms of the potential.

DEFECT ENERGY

A molecular dynamics simulation to determine the formation energy of a vacancy and a interstitial in copper crystals was performed by using above EAM potentials. A model crystal consisted of 1300 atoms was arranged, and a vacancy or a interstitial was introduced near the center of the crystal. The time interval Δt of the MD step is chosen to be 2×10^{-15} sec. After 10000 MD steps, the system attain an equilibrium state. The total potential energy is compared with that of the perfect crystal containing same number of atoms, and the defect
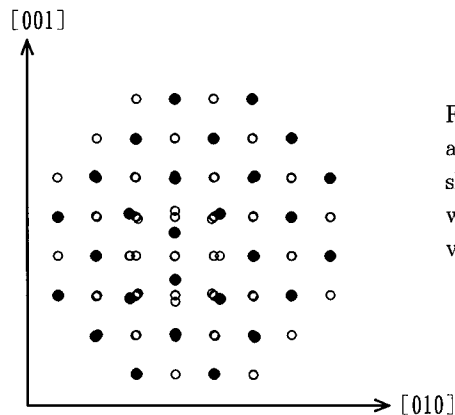
Fig.2 Atomic configuration around a [100]-split interstitial. Solid circles show the atoms on a [100]-plane in which the interstitial atoms are involved.

Table III   Formation energy of vacancy E_V and interstitial E_I in copper (in units of eV)

	A-T-V-F	O-J	D-K	Exp.
E_V	1.05	1.23	1.04	1.30
E_I	4.31	3.48	3.76	3.94

energy is determined. As an example, atomic configuration around an interstitial is shown in Fig.2. The split interstitial in [100] direction is found to be a stable configuration. The determined formation energies of the vacancy and the interstitial by using the three EAM potentials are summarized in Table III. The values determined from the three EAM potentials are not so different as seen in the calculation of the Grüneisen parameter. The anharmonicity is considered to be less sensitive to the defect formation energy.

SUMMARY

The empirical interatomic potential has conveniently been used in the molecular dynamics simulations, especially for the large atomic systems. The typical EAM potential functions have been characterized through the calculation of the anharmonicity and of the defect energy. The Grüneisen parameter is found to depend on the potential function , strongly. The formation energies of the vacancy and the interstitial also depend on the potential but less than the Grüneisen parameter. The characterization of the potential on the points of lattice vibration and the glass transition is now in progress.

ACKNOWLEDGMENT

REFERENCES

1. M.S. Daw and M.I. Baskes; Phys. Rev. B, 29 (1984) 6443.
2. M.W. Finnis and J.E. Sinclair; Phil. Mag. A, 50 (1984) 45.
3. G. J. Ackland, G. Tichy, V. Vitek M. W. Finnis; Phil. Mag. A, 56 (1987) 735.
4. D.J. Oh and R.A. Johnson; J. Mater. Sci., 3 (1988) 471.
5. M. Doyama and Y. Kogure; Radiation Effects and Defects in Solids, 142 (1997) 565.
6. K. A. Gschneidner, Jr; Solid State Physics, 16 (1964) 275.

A TIGHT-BINDING MODEL FOR OPTICAL PROPERTIES OF POROUS SILICON

M. CRUZ[†], M.R. BELTRAN, C. WANG, and J. TAGÜEÑA-MARTINEZ[‡],

Instituto de Investigaciones en Materiales, UNAM,

Apdo. Postal 70-360, 04510, D.F. México.

[†] Escuela Superior de Ingeniería Mecánica y Eléctrica - UC, IPN, México.

[‡]Centro de Investigación en Energía, UNAM,

A.P. 34, C.P. 62580, Temixco, Mor., México.

ABSTRACT

Semi-empirical tight-binding techniques have been extensively used during the last six decades to study local and extended defects as well as aperiodic systems. In this work we propose a tight-binding model capable of describing optical properties of disordered porous materials in a novel way. Besides discussing the details of this approach, we apply it to study porous silicon (p-Si). For this purpose, we use an sp^3s^* basis set and supercells, where empty columns are digged in the [001] direction in crystalline silicon (c-Si). The disorder of the pores is considered through a random perturbative potential, which relaxes the wave vector selection rule, resulting in a significant enlargement of the optically active k-zone. The dielectric function and the light absorption spectra are calculated. The results are compared with experimental data showing a good agreement.

INTRODUCTION

From a theoretical point of view, the study of optical properties starts with the calculation of the electronic band structure. This can be achieved mainly by two possible approaches. First principle methods are very successful in treating small systems. However, semi-empirical or tight-binding calculations, less computationally intensive than first principle calculations, can consequently deal with larger and more complex structures. The use of phenomenological parameters in semi-empirical calculations includes many-body effects, otherwise extremely difficult to consider in first principle methods. Another advantage of the tight-binding approach is the possibility of treating different types of disorder such as local defects, alloys, quasicrystals and amorphous systems. Several well known techniques to treat disorder exist in the literature such as virtual crystal approximation (VCA), average T-matrix approximation (ATA) and coherent potential approximation (CPA). These methods have played an important rôle in the understanding of many new materials though they are mostly appropriate for local disorder and, if self-consistency is included they become extremely computing demanding.

We present here a semi-empirical tight-binding supercell model [1] capable to study disordered and nano structured porous media. Our aim is to calculate optical properties from

the electronic band structure. In order to address the disorder, instead of using the conventional procedures, we introduce a new approach which includes a random perturbative potential to simulate the aleatory distribution of pores. This random potential produces a relaxation of the **k** wavevector selection rule broadening the optically active zone.

One of the dominant trends in current research of materials science and solid state physics is the study of nanometric materials and devices. In these nanostructures many new effects take place and it is believed that the quantum confinement plays a mayor rôle in their peculiar behaviours. In particular, porous silicon (p-Si) represents a very interesting, easily produced and non-expensive nanostructured system.

Although the supercell model could be applied to different porous systems and nanostructures, in this paper we describe its application to porous silicon, which is produced from crystalline silicon by means of an electrochemical etching [2]. This process leaves a porous or sponge-like structure, with diameters at nanometer scale and lengths at micron scale, which intermingle at random. This material possess both, order and disorder, but at separate length scales. The order lies in the crystalline structures, and the disorder originates from the different shapes, orientations and distribution of the crystalline silicon crystallites, which conform the randomly arranged branches, characteristic of p-Si. The enormous interest in this material is mainly due to its very efficient luminescence at room temperature, which could be applied to all silicon based optoelectronic technology. For a detailed and recent review of p-Si see, for example, the Ref. [3].

From the theoretical point of view, calculations have primarily been performed on idealized Si wire and dot structures [4]. Surface dangling bonds are passivated with hydrogen, as is in general the case in freshly etched Si. First principle calculations on p-Si have been performed using density function theory (DFT) within the local density approximation (LDA) [5, 6] which have been demonstrated to be successful in dealing with ground state properties of Si. Semi-empirical calculations of band structures of idealized p-Si structures have been performed [1, 4, 7, 8], giving band gaps in agreement with the experimental trends.

On the experimental side, extensive studies have been achieved [9]. In particular, we are interested on optical properties of p-Si, such as the dielectric function, measured by Koshida et. al [10] amongst others, and the absorption coefficient [11, 12].

In the next section we describe the supercell model with disorder. Following this, we present and discuss the results obtained. Finally, some conclusions are given.

THEORY

We use a minimum basis capable of describing an indirect band gap along the X-direction, an sp^3s^*, with P. Vogl, H.P. Hjalmarson and J. Dow's parameters [13], which gives an 1.1 eV gap in bulk crystalline silicon. Empty columns (pores) are produced by removing, in the [001] direction, columns of atoms within supercells. We saturate the pore surface with hydrogen atoms. The Si-H bond length is taken as 1.48 angstroms. The on-site energy of the H atom is considered to be -4.2 eV, since the free H atom energy level, -13.6 eV, is so close to the s-state energy level of a free Si atom, -13.55 eV [14], therefore the on-site energy of H is taken to be the same as that of silicon, as in [15]. The H-Si orbital interaction parameters are taken as $ss\sigma_{H-Si} = -4.075$ eV, $sp\sigma_{H-Si} = 4.00$ eV, which are obtained by fitting the energy levels of silane [16].

The real (ϵ_1) and the imaginary (ϵ_2) parts of the dielectric function ($\epsilon = \epsilon_1 + i\epsilon_2$) are related by the Kramers-Kronig dispersion relations and ϵ_2 can be calculated as [14],

$$\epsilon_2(\omega) \propto \frac{1}{\omega^2} \sum_{k_v,k_c} |\langle \mathbf{k}_v | \mathbf{p} \cdot \mathbf{e} | \mathbf{k}_c \rangle|^2 \delta(E(\mathbf{k}_c) - E(\mathbf{k}_v) - \hbar\omega), \tag{1}$$

where $\mathbf{e}$ is the light polarization vector, $\mathbf{p}$ is the electronic linear momentum, $|\mathbf{k}_v\rangle$ and $|\mathbf{k}_c\rangle$ are valence- and conduction-band eigenstates, respectively. In the tight-binding scheme eigenstates $|\mathbf{k}\rangle = \sum_{i,\mu} a_{i,\mu}^{\mathbf{k}} |i,\mu\rangle$, where i is the site index and μ identifies the orbital, then the dipole matrix in the equation (1) can be expressed as:

$$\langle \mathbf{k}_v | \mathbf{p} \cdot \mathbf{e} | \mathbf{k}_c \rangle = \sum_{i,j,\mu,\nu} a_{i,\mu}^{v}{}^* a_{j,\nu}^{c} \langle i\mu | \mathbf{p} | j\nu \rangle \cdot \mathbf{e}. \tag{2}$$

The dipole matrix elements in Eq. (2) may be rewritten in terms of the Hamiltonian (H) and the position ($\mathbf{r}$) operators, using the commutation relation $\mathbf{p} = \frac{im}{\hbar}[H, \mathbf{r}]$,

$$\langle i\mu | \mathbf{p} | j\nu \rangle = \frac{im}{\hbar} \sum_{l,\lambda} (\langle i\mu | H | l\lambda \rangle \langle l\lambda | \mathbf{r} | j\nu \rangle - \langle i\mu | \mathbf{r} | l\lambda \rangle \langle l\lambda | H | j\nu \rangle). \tag{3}$$

Since the polarizability of a free atom is much smaller than that of the corresponding semiconductor [17], Eq. (3) can be simplified as [18]:

$$\langle i\mu | \mathbf{p} | j\nu \rangle = \frac{im}{\hbar} \langle i\mu | H | j\nu \rangle \mathbf{d_{i,j}}, \tag{4}$$

where $\mathbf{d_{i,j}} = \langle j\nu | \mathbf{r} | j\nu \rangle - \langle i\mu | \mathbf{r} | i\mu \rangle$ is the distance between the gravity centers of the orbitals μ and ν placed at atoms i and j, respectively, and it is independent on orbitals if the crystal field is symmetric. Notice that the contribution to the dipole matrix coming from two orbitals at the same atom is neglected.

In order to understand the optical properties of p-Si, several unique features inherent to this material have to be considered. The luminescence of p-Si depends largely on its skeleton size, its disordered nature, as well as on its surface chemical composition. These characteristics switch on nonmomentum-conserving, no-phonon assisted radiative transitions [3]. In this work, we propose to include a random perturbative potential in the Hamiltonian to extend the supercell model, representing the aleatory distribution of pores. Its Fourier transform introduces an additional $\mathbf{k}$ wavevector to be considered in the momentum conservation rule. Consequently, due to the randomness, when pores are introduced, we take into account both, vertical and non-vertical radiative interband transitions. However, not all of them contribute equally, depending on the nature of the perturbative potential considered. In principle, it is possible to estimate the interval of non-vertical $\mathbf{k}$ values, defining a window around each vertical transition. An analysis of the behaviour of the $\mathbf{k}$-window is given in the next section.

RESULTS

To calculate optical properties of p-Si, we have considered interband transitions for an x-direction polarized light. They are calculated in 8-atom supercells [1] with a 1-atom columnar pore saturated by hydrogen atoms. The calculation has been performed by
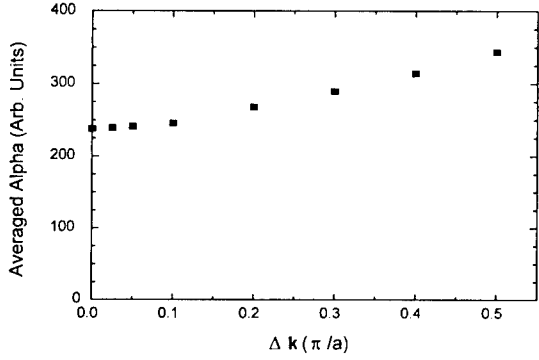
Figure 1: Analysis of the averaged absorption coefficient ($\overline{\alpha}$) as a function of the window width ($\Delta \mathbf{k}$) introduced due to non-vertical transitions between valence and conduction band states, and $a = 5.43$ angstroms is the supercell size.

considering 1331 $\mathbf{k}_c$-points in the first Brillouin zone. For each $\mathbf{k}_c$ point, a corresponding cubic $\mathbf{k}_v$-window of different sizes is considered.

We start by analyzing the rôle of the disorder in p-Si. Figure 1 shows the thermally averaged absorption coefficient ($\overline{\alpha}$) as a function of the $\mathbf{k}$-space window width. The absorption coefficient ($\alpha(\omega)$) is defined as

$$\alpha(\omega) = \frac{\omega}{nc}\epsilon_2(\omega)$$

where n is the refraction index of the porous media and c is the light speed. The thermally averaged absorption coefficient, proportional to the recombination rate, is given by

$$\overline{\alpha} = \frac{\int \alpha(\omega)\exp(-\frac{\hbar\omega}{k_B T})d\omega}{\int \exp(-\frac{\hbar\omega}{k_B T})d\omega}.$$

Notice that the absorption increases as the window broadens, which is expected since more non-vertical transitions are considered. It can also be observed that the order of magnitude of the absorption is maintained, in spite that a significant fraction of the Brillouin zone is included in the window. It is worth mentioning that this result is much larger than that obtained if the same calculation is performed on crystalline silicon. This means that the quantum confinement plays a substantial rôle in the optical properties of p-Si.

In figure 2 the results of the absorption coefficient, including non-vertical transitions, are compared with experimental data obtained by Sagnes *et al* [11] for a p-type 74% porosity sample. We have shifted the absorption curve to fit the experimental enery gap in order to analyze its shape, since the theoretical and the reported porosities are not equivalent. Notice that there is a fair agreement between theory and experiment around the band gap edge, in spite that no d-orbitals neither other elementary excitations have been considered
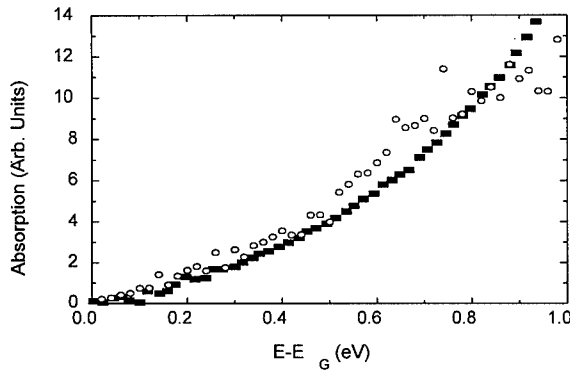
Figure 2: Absorption coefficient $\sqrt{\alpha E}$ (open circles) calculated from an 8-atom supercell with a one-atom columnar pore, compared with Sagnes' experimental data (solid squares) for a p-type porous silicon sample with 74% porosity.

In general, the analysis of the absorption coefficient around the band gap could reveal much about the gap nature, and figure 2 suggests that disorder is important. In a recent joint density of states study [19] considering 3D disordered crystallites they also reproduce the same absorption data and conclude that disorder plays a very important role.

CONCLUSIONS

We have shown previously [20] that a simple quantum mechanical treatment, such as a phenomenological tight-binding technique, is capable of reproducing the essential features of the dielectric function of p-Si. Furthermore, from the comparison with the experimental data we can conclude that the tight-binding supercell model, when disorder is considered, gives the correct shape of the absorption coefficient around the band gap in p-Si.

Certainly, porous silicon is a complex material, where the quantum confinement and the disorder are two of the most important features. The quantum confinement effect is considered within the supercell method. However, disorder is difficult to treat within the widely used theoretical techniques, such as CPA. The non-vertical transition approach discussed here provides us a non-conventional first-approximation tool to treat the disorder present in p-Si, that has the advantage of being simple to implement and gives reasonably good results.

This calculation is currently being extended to calculate the radiative lifetime of excited states. On the other hand, further studies should be performed to quantify the weight of each non-vertical transition, which will depend on the specific random potential to be included in the Hamiltonian.

References

[1] M. Cruz, C. Wang, M.R. Beltrán and J. Tagüeña-Martínez, Phys. Rev. B **53**, p. 3827 (1996).

[2] L.T. Canham, Appl. Phys. Lett. **57**, p. 1046 (1990).

[3] A.G. Cullis, L.T. Canham and P.D.J. Calcott, Appl. Phys. Rev. **82**, p. 909, (1997).

[4] C. Delerue, G. Allan, and M. Lannoo, Phys. Rev. B **48**, p. 11024 (1993); J. P. Proot, C. Delerue, and G. Allan, Appl. Phys. Lett. **61**, p. 1948 (1992).

[5] F. Buda, J. Kohanoff, and M. Parrinello, Phys. Rev. Lett. **69**, p. 1272 (1992).

[6] B. Delley, and E.F. Steigmeier, Appl. Phys. Lett.,**67**, p. 2370 (1995).

[7] C. Delerue, M. Lannoo, G. Allan, N.A. Hill and K.B. Whaley, Phys. Rev. Lett. **76**, p. 3038 (1996).

[8] G.D. Sanders and Y.C. Chang, Phys. Rev. B **45**, p. 9202 (1992).

[9] L. Tsybeskov, et al., Light Emitting Porous Silicon : Materials Science Properties and Device Applications, IEEE, J. of Special Topics in Quantum Elect. **1**, pp. 1126 (1995).

[10] N. Koshida, et al., Appl. Phys. Lett. **63**, p. 2774 (1993).

[11] I. Sagnes, et al., Appl. Phys. Lett. **62**, p. 1155 (1993).

[12] L. Tsybeskov, S.P. Duttagupa, K. D. Hirschman and P. Fauchet, Appl. Phys. Lett. **68**, p. 2058 (1996).

[13] P. Vogl, H.P. Hjalmarson, and J.D. Dow, J. Phys. Chem. Solids **44**, p. 365 (1983).

[14] W.A. Harrison, Electronic Structure and the Properties of Solids (Dover, NY, 1989).

[15] Shang Yuan Ren and John D. Dow, Phys. Rev. B **45**, p. 6492 (1992).

[16] Fu Huaxiang, Ye Ling, and Xide Xie, Phys. Rev. B **48**, p. 10978 (1993).

[17] L. Brey and C. Tejedor, Solid State Commun. **48**, p. 403 (1983).

[18] B. Koiller, R. Osório, and L.M. Falicov, Phys. Rev. B **43**, p. 4170 (1991).

[19] P.M. Derlet, T.C: Choy, A.M. Stoneham, J. Phys.: Condens. Matter **7**, p. 2507 (1995).

[20] M. Cruz, et al., Mat. Res. Soc. Symp. Proc. **452**, p. 69 (1997).

TIGHT-BINDING ELECTRON-ION DYNAMICS:
A METHOD FOR TREATING NONADIABATIC PROCESSES
AND INTERACTIONS WITH ELECTROMAGNETIC RADIATION

J.S. GRAVES, R.E. ALLEN
Department of Physics, Texas A&M University, College Station, Texas, 77843

ABSTRACT

A method is introduced for simulations of the coupled dynamics of electrons and ions in a molecule or material. It is applicable to general nonadiabatic processes, including interactions with an arbitrarily intense radiation field. The field is included in the electronic Hamiltonian through a time-dependent Peierls substitution. The time-dependent Schrödinger equation is solved with an algorithm that preserves orthogonality, and the atomic forces are obtained from a generalized Ehrenfest theorem. Calculations for GaAs and Si demonstrate that the method is reliable and quantitative.

Many processes in physics, chemistry [1-3], and biology [4-6] involve the interaction of electromagnetic radiation with complex molecules and materials. Traditional treatments of this problem involve the Born-Oppenheimer approximation (in which the electrons are assumed to adiabatically follow the motion of the nuclei), the Franck-Condon principle (in which the nuclei are regarded as frozen during each electronic transition), and Fermi's golden rule (which is based on both first-order perturbation theory and the premise that the field varies harmonically on a long time scale). These assumptions may be difficult to employ for a complex system, and they are not always valid. For example, Fermi's golden rule is invalid for ultra-intense and ultrashort laser pulses [7-9]. In the most general case, one needs numerical simulations.

Here we introduce a method for simulations of the coupled dynamics of valence electrons and ion cores in an arbitrarily strong and time-dependent electromagnetic field. Although a first-principles formulation is possible [10], a tight-binding representation is preferable for practical calculations: (1) The electronic excitations play a central role, so it is important that the excited states be at their proper energies. (These are fitted to experiment in a semiempirical tight-binding model, whereas they are typically too low in the local density approximation and too high in Hartree-Fock.) (2) Since the time step is of order 1 femtosecond, and the system may contain many atoms, the method must be computationally fast. (3) A tight-binding representation involves chemically-meaningful basis states which are localized on the atoms, and which has the same symmetries as atomic orbitals. One can then immediately interpret the results using intuitive ideas based on ground-state and excited-state chemistry [11].

Tight-binding electron-ion dynamics begins with the model Lagrangian [10]

$$L = \sum_{l\kappa} \frac{1}{2} M_l \dot{X}_{l\kappa}^2 - U_{rep} + \sum_j \mathbf{\Psi}_j^\dagger \cdot \left(i\hbar \frac{\partial}{\partial t} - \mathbf{H} \right) \cdot \mathbf{\Psi}_j. \tag{1}$$

The first term is the kinetic energy of the ions, with coordinates $X_{l\kappa}$, which are treated classically. The second is a summation over repulsive potentials which model the ion-ion repulsion,

together with the negative of the electron-electron repulsion which is doubly counted in the third term [12]. This last term is the tight-binding version of the standard Lagrangian for particles treated in a time-dependent self-consistent-field approximation [13]. Each electron, labeled by j, has its own time-dependent state vector $\boldsymbol{\Psi}_j$. If there are N tight-binding basis functions in the system, $\boldsymbol{\Psi}_j$ is N-dimensional, and the time-dependent Hamiltonian $\mathbf{H}$ is $N \times N$.

When the action $\int L dt$ is required to be stationary with respect to variations in $X_{l\kappa}$ and $\boldsymbol{\Psi}_j^\dagger$, we obtain the equations of motion for electron and ions:

$$i\hbar \partial \boldsymbol{\Psi}_j / \partial t = \mathbf{H}(t) \boldsymbol{\Psi}_j \tag{2}$$

$$M\ddot{X} = -\sum_j \boldsymbol{\Psi}_j^\dagger \cdot \frac{\partial \mathbf{H}}{\partial X} \cdot \boldsymbol{\Psi}_j - \frac{\partial U_{rep}}{\partial X} \tag{3}$$

where X is any ion coordinate. These are respectively the time-dependent Schrödinger equation and Ehrenfest's theorem, with the electrons treated in a tight-binding picture and the ions treated classically. The corresponding equations in a first-principles formulation were obtained previously, using a completely different derivation.

The electrons and ions are coupled in (2) and (3), because $\mathbf{H}$ is a function of the ion coordinates and the forces on the ions are influenced by the electronic states. We now need to couple the electrons to the radiation field. (One can also easily couple the ions to the electromagnetic field, but this is a minor effect if the field oscillates on a one femtosecond time scale, two orders of magnitude smaller than the response time of the ions.) The most convenient way to introduce the field into the electronic Hamiltonian is to employ a time-dependent Peierls substitution [14]: First consider the standard Hamiltonian in the coordinate representation with a time-dependent electromagnetic vector potential $\mathbf{A}$:

$$H = (\mathbf{p} - \frac{e}{c}\mathbf{A})^2 / 2m + V(\mathbf{r}) \tag{4}$$

where $\mathbf{p} = -i\hbar\nabla$ and e is negative for an electron. This is equivalent to

$$H = \exp\left(-\frac{ie}{\hbar c} \int \mathbf{A} \cdot d\mathbf{r}\right) H^0 \exp\left(\frac{ie}{\hbar c} \int \mathbf{A} \cdot d\mathbf{r}\right) \tag{5}$$

with

$$H^0 = \mathbf{p}^2 / 2m + V(\mathbf{r}) \tag{6}$$

as one can easily verify by substituting (6) into (5) and letting H operate on an arbitrary function $\Psi(\mathbf{r})$. To employ (5) in a tight-binding scheme, we recognize that the matrix elements of (5) are the same as matrix elements of (6) with the localized basis functions $\phi_\alpha(\mathbf{r} - \mathbf{R})$ multiplied by $\exp(ie \int \mathbf{A} \cdot d\mathbf{r}/\hbar c)$. In this factor, it is consistent with the spirit of tight-binding to take $\int \mathbf{A} \cdot d\mathbf{r} \approx \mathbf{A} \cdot \mathbf{r} \approx \mathbf{A} \cdot \mathbf{R}$, provided that $\mathbf{A}$ is slowly-varying on an atomic scale (as it is for electromagnetic radiation with $\hbar\omega \lesssim 10$ eV). Then the matrix elements $H_{\alpha\beta}(\mathbf{R} - \mathbf{R}')$ are modified by the Peierls substitution

$$H_{\alpha\beta}(\mathbf{R} - \mathbf{R}') = H^0_{\alpha\beta}(\mathbf{R} - \mathbf{R}') \exp\left(\frac{ie}{\hbar c}\mathbf{A} \cdot (\mathbf{R} - \mathbf{R}')\right) \tag{7}$$

This approach requires no additional parameters and is valid for arbitrarily strong time-dependent fields.

Solution of the ionic equations of motion (3) is essentially the same as in tight-binding molecular dynamics [12,15,16]. (The velocity Verlet algorithm was used.) Solution of (2) requires more care, since a naive algorithm for this first order equation fails to conserve probability. Instead we write the time-evolution equation in the form

$$\exp\left(\frac{i\mathbf{H}\Delta t}{2\hbar}\right) \cdot \boldsymbol{\Psi}(t + \Delta t) = \exp\left(\frac{-i\mathbf{H}\Delta t}{2\hbar}\right) \cdot \boldsymbol{\Psi}(t). \tag{8}$$

Now approximate the exponential by its first two terms. In lowest order this gives the Cayley algorithm

$$\boldsymbol{\Psi}(t + \Delta t) = (1 + \frac{i\mathbf{H}\Delta t}{2\hbar})^{-1} \cdot (1 - \frac{i\mathbf{H}\Delta t}{2\hbar}) \cdot \boldsymbol{\Psi}(t). \tag{9}$$

With a time step Δt of 50 attoseconds, we then find that probability is conserved to within about 3 parts in 10^6 during the full simulation, and orthogonality is preserved at this same level. Energy is conserved to about one part in 10^6 with no applied field, or one part in 10^4 after interaction with an intense laser pulse (3×10^{12} W/cm^2 for about 100 femtoseconds).

Using the method outlined above, we have performed calculations for the electronic and structural response of semiconductors to ultra-intense and ultrashort laser pulses [17-20]. A standard tight-binding model was used [21], together with a standard r^{-2} scaling for the interatomic matrix elements [22]. We used a nonstandard repulsive potential with the form

$$\phi(r) = \frac{\alpha}{r^4} + \frac{\beta}{r^6} + \frac{\gamma}{r^8}. \tag{10}$$

The three parameters α, β, and γ were fitted to the experimental values of the cohesive energy, interatomic spacing, and bulk modulus. A cubical cell containing eight atoms was used in the simulations, and periodic boundary condition were imposed on the motion of the ions. The electronic states are then Bloch states corresponding to this large unit cell.
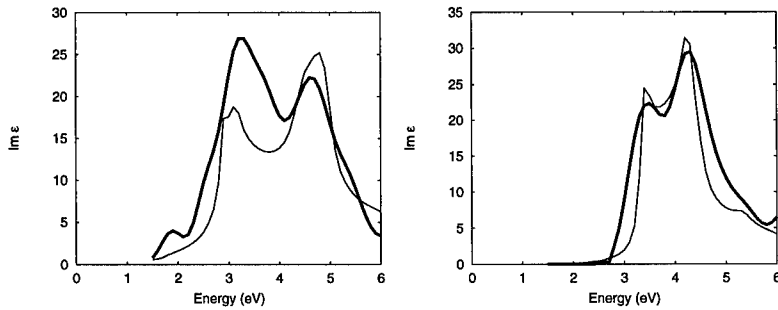
Fig. 1. Im $\varepsilon(\omega)$ calculated in the present work (heavy curves), compared with the measurements of Ref. 23 (light curves). Left panel: GaAs. Right panel: Si.

The time dependence of the electronic states and ionic positions was calculated as described above, and the imaginary part of the dielectric function was obtained from the formula

$$\mathrm{Im}\ \epsilon(\omega) \quad \propto \quad \frac{1}{\omega^2} \sum_{n,m,\mathbf{k}} [f_n(\mathbf{k}) - f_m(\mathbf{k})]\mathbf{p}_{nm}(\mathbf{k}) \cdot \mathbf{p}_{mn}(\mathbf{k})\delta(\omega - \omega_{mn}(\mathbf{k})) \tag{11}$$

where the notation is defined in Ref. 14. This result of linear response theory is valid in the probe phase of a pump-probe experiment. It also yields good agreement with more usual optical measurements [23], as illustrated in Fig. 1 for GaAs and Si (or in Ref. 14 for GaAs with a slightly different tight-binding model). In Fig. 2, Im $\varepsilon(\omega)$ was calculated in arbitrary units, and the height of the theoretical curve then adjusted for a better comparison with the measurements. Without this adjustment, theory and experiment still agree to within a factor of two [14]. A set of 512 sample k-points were used in calculating the time-dependent dielectric function during the simulations. Further details are given elsewhere [24].
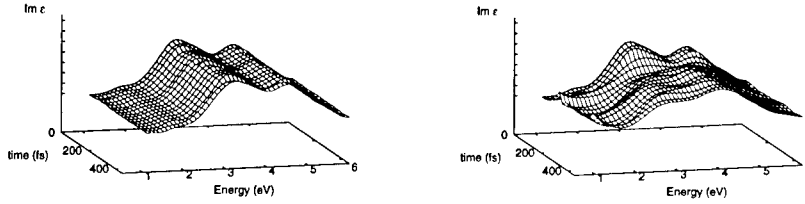
Fig. 2. Time-dependent dielectric function (in arbitrary units) for GaAs, with $\hbar\omega = 1.95$ eV and a FWHM pulse duration of 70 fs. Left panel: fluence = 0.815 kJ/m^2. Right panel: fluence = 3.26 kJ/m^2.
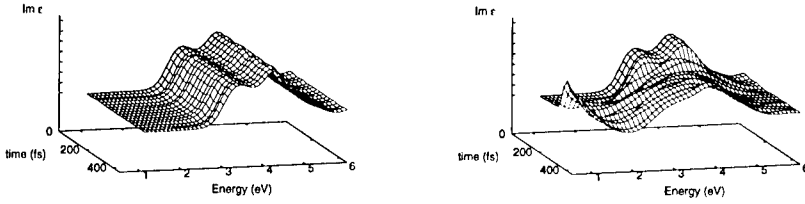
Fig. 3. Time-dependent dielectric function for Si, with $\hbar\omega = 1.95$ eV and a FWHM pulse duration of 70 fs. Left panel: fluence = 0.815 kJ/m^2. Right panel: fluence = 3.26 kJ/m^2.

Representative results for the time dependence of $\varepsilon(\omega)$, during and after the application of a laser pulse, are shown in Fig. 2 for GaAs and Fig. 3 for Si. In each case, notice the difference in the behavior of $\varepsilon(\omega)$ for pulses of lower and higher intensity. For the pulse of lower intensity, Im $\varepsilon(\omega)$ is zero at all times for $\hbar\omega$ less than the band gap, demonstrating that the material remains a semiconductor and there is no absorption within this range of energies. In addition, the structural features of Fig. 1 persist, showing that the original bandstructure remains intact. For the pulse of higher intensity, on the other hand, below-bandgap absorption is observed soon after the pulse, indicating a transition to metallic behavior. The original structural features in $\varepsilon(\omega)$ are also washed out following the pulse. This general behavior is consistent with that observed in the experiments [17,18]. A more detailed analysis of the lattice destabilization, band-gap collapse, and modification of the dielectric function is presented elsewhere [24].
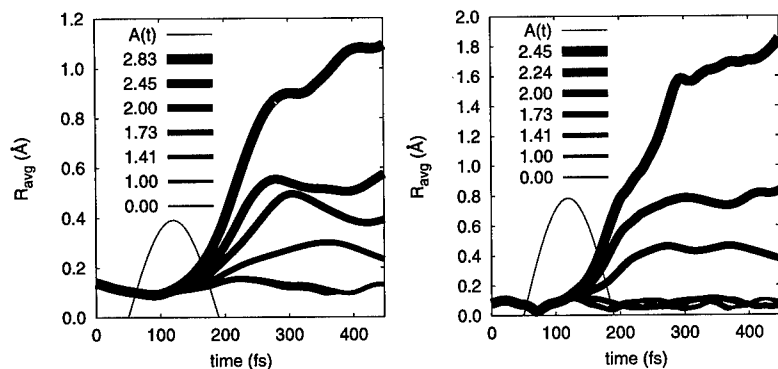
Fig. 4. Average distance moved by an atom, during and following a laser pulse. The amplitude A_0 of the vector potential is given in gauss·cm. Left panel: GaAs. Right panel: Si.

In summary, we have introduced a method for simulations of the coupled dynamics of electron and ions in a molecule or material which is subjected to arbitrarily intense electromagnetic radiation. We have also applied this method in studies of GaAs and Si subjected to ultra-intense and ultrashort laser pulses. The results for the time dependence of the dielectric function $\varepsilon(\omega)$ are in good agreement with what is observed in pump-probe experiments on these materials [17,18]. We find that the modification of $\varepsilon(\omega)$ indeed results from a non-thermal lattice destabilization and band-gap collapse, on a 100 femtosecond time scale, when electrons are promoted from bonding valence-band states to antibonding conduction-band states.

ACKNOWLEDGEMENT

REFERENCES

1. G. J. Ferraudi, *Elements of Inorganic Photochemistry* (Wiley, New York, 1988).

2. F. A. Casey and R. J. Sundberg, *Advanced Organic Chemistry* (Plenum, New York, 1993).

3. H.-L. Dai and W. Ho, *Laser Spectroscopy and Photochemistry on Metal Surfaces* (World Scientific, Singapore, 1995), Vols. 1 and 2.

4. E. Kohen, R. Santus, and J. G. Hirschberg, *Photobiology* (Academic, Boston, 1995).

5. C. K. Mathews and K. E. Van Holde, *Biochemistry*, second edition (Benjamin/ Cummings, Menlo Park, 1990).

6. N. A. Campbell, *Biology*, second edition (Benjamin/ Cummings, Menlo Park, 1990).

7. C. J. Joshi and P. B. Corkum, Physics Today **48**, 36 (1995).

8. *Atoms in Intense Laser Fields*, edited by M. Gommila (Academic, Boston, 1992).

9. *Molecules in Laser Fields*, edited by A. D. Bandrank (Marcel Dekker, New York, 1994).

10. R. E. Allen, Phys. Rev. B **50**, 18629 (1994).

11. *The Chemical Bond, Structure and Dynamics*, edited by A. Zewail (Academic, Boston, 1992).

12. M. Menon and R. E. Allen, Phys. Rev. B **33**, 7099 (1986).

13. A. de Shalit and H. Feshbach, *Theoretical Nuclear Physics* (Wiley, New York, 1974), Vol. 1, p. 530.

14. M. Graf and P. Vogl, Phys. Rev. B **51**, 4940 (1995).

15. O. F. Sankey and R. E. Allen, Phys. Rev. B **33**, 7164 (1986).

16. M. Sawtarie, M. Menon, and K. R. Subbaswammy, Phys. Rev. B **49**, 7739 (1994), and references therein.

17. E. N. Glezer, Y. Siegal, L. Huang, and E. Mazur, Phys. Rev. B **51**, 6959,9589 (1995).

18. K. Sokolowski-Tinten, J. Bialkowski, and D. von der Linde, Phys. Rev. B **51**, 14186 (1995).

19. S. V. Govorkov, Th. Schröder, I. L. Shumay, and P. Heist, Phys. Rev. B **46**, 6864 (1992).

20. C. V. Shank, R. Yen, and C. Hirlimann, Phys. Rev. Lett. **50**, 454 (1983).

21. P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

22. W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).

23. D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).

24. J. S. Graves, Ph.D. thesis, Texas A&M University, 1997.

OPTICAL PROPERTIES OF MATERIALS USING THE EMPIRICAL TIGHT-BINDING METHOD

L. C. LEW YAN VOON
Department of Physics, Worcester Polytechnic Institute, 100 Institute Road, Worcester, Massachusetts 01609

ABSTRACT

Procedures for calculating the optical properties of periodic systems using the empirical tight-binding method are compared. Results of the linear and nonlinear susceptibilities are presented using a recently developed exact formalism [Lew Yan Voon and Ram-Mohan, Phys. Rev. B**47**, 15500 (1993)].

INTRODUCTION

The multiband empirical tight-binding method was developed by Slater and Koster [1] in 1954 in order to provide an efficient full Brillouin zone parametrization of the electronic band structures of periodic systems. The Hamiltonian was formally represented in terms of a finite number of localized atomiclike basis functions satisfying the Bloch condition, the functions themselves never being explicitly given. The unknown parameters (the tight-binding [TB] parameters) were then obtained by fitting the eigenvalues of the Hamiltonian matrix (i.e., energies) to experimental data. For the next 30 years or so, the TB model was used extensively to solve a wide variety of problems: point defects levels, surface and heterostructure band structures, bulk and surface structures, transport coefficients. Missing was the knowledge of how to calculate optical coefficients (e.g., absorption coefficient) exactly and completely. The argument was that the calculation of the optical coefficients requires first calculating momentum matrix elements and the latter could not be done because the actual explicit wave functions were never known. Approximate schemes were introduced which relied mainly on fitting the momentum operator in a fashion analogous to the Hamiltonian [2, 3]. In 1993, we managed to prove explicitly that the momentum operator could be defined exactly in terms of the Hamiltonian operator [4].

Our aim here is to compare the fitting methods to our exact expression. We also present results of the linear and nonlinear optical susceptibilities obtained using our procedure to illustrate the difference between different TB parameter sets for optical properties.

THEORY

The long-wavelength optics we are interested in here is obtained from coefficients characteristic of the material. For linear optics, we will refer to the imaginary part of the dielectric function,

$$\epsilon_2(\omega) = \frac{\pi e^2}{\epsilon_0 m_0^2 \omega^2 V} \sum_{cv\mathbf{k}} f_v(E_v) \left[1 - f_c(E_c)\right] |\langle c|p|v \rangle|^2 \delta(E_{cv} - \hbar\omega), \tag{1}$$

while, for second-order nonlinear optics, we will compute the coefficient of second-harmonic

generation [5]:

$$\chi_{ijk}^{(2)}(2E) = \frac{ie^3\hbar^3}{2\epsilon_0 m_0^3 E^3} \times \sum_{vcc'\mathbf{k}} f_v \left\{ \frac{p_{c'v}^i \{p_{vc}^j p_{cc'}^k\}}{(E_{c'v} + 2E)(E_{cv} + E)} + \cdots \right\}. \tag{2}$$

For the above, one needs the band structure $E_{n\mathbf{k}}$ and the momentum matrix element $\mathbf{p}_{nm}(\mathbf{k})$.

In computing the momentum matrix element within the TB formalism, a number of procedures had been adopted in the past. The most appealing was to represent the momentum operator in the same basis set as the Hamiltonian:

$$\mathbf{p}_{nm}(\mathbf{k}) = \langle n\mathbf{k}|\hat{\mathbf{p}}|m\mathbf{k}\rangle = \sum_{\alpha\beta bb'} C_{\alpha b}^*(n\mathbf{k})C_{\beta b'}(m\mathbf{k}) \sum_l e^{i\mathbf{k}\cdot\mathbf{R}_{lbb'}}$$

$$\times \int d\tau\, \phi_{\alpha b}^*(\mathbf{r})\hat{\mathbf{p}}\phi_{\beta b'}(\mathbf{r} - \mathbf{R}_{lbb'}) \tag{3}$$

where

$$|n\mathbf{k}\rangle = \frac{1}{N} \sum_{\alpha b} C_{\alpha b}(n\mathbf{k}) \sum_l e^{i\mathbf{k}\cdot\mathbf{R}_{lb}} \phi_\alpha(\mathbf{r} - \mathbf{R}_{lb}), \tag{4}$$

is the expansion of the band states in terms of the TB Bloch states for orbitals ϕ_α located at position $\mathbf{R}_{lb}$. Symmetry is then used to reduce the momentum matrix elements into a minimal set of independent quantities. Thus, Xu [2] assumed that intraatomic excitations dominate and rewrote the expansion as

$$\mathbf{p}_{nm}(\mathbf{k}) = \sum_{\alpha\beta b} C_{\alpha b}^*(n\mathbf{k})C_{\beta b}(m\mathbf{k})\mathbf{P}_{\alpha\beta}^b. \tag{5}$$

The unknown quantities $\mathbf{P}_{\alpha\beta}^b$ were then obtained by fitting to experimental dielectric functions. Chang and Aspnes [3] used a similar procedure but obtained a slightly different set of parameters, which they then fitted to empirical pseudopotential calculations. Explicit values for GaAs are given in Table 1. P_{aa}, for example, refers to the intraatomic momentum

Table 1: Values of the momentum matrix elements for GaAs obtained by Chang and Aspnes [3].

	P_{aa}	P_{cc}	P_{ss}	P_{sp}	P_{ps}	$\ldots$
GaAs (eV)$^{1/2}$	6.111	3.023	0.044	-0.114	0.178	$\ldots$

matrix element between an s orbital and a p orbital for the anion. P_{sp} is a nearest-neighbor momentum matrix element. These numerical values are consistent with the model of Xu since the intraatomic momentum matrix elements are much bigger than the nearest-neighbor ones. The procedures outlined above have a number of flaws. If the system is under strain, momentum matrix elements would have to be obtained for different levels of strain. The models are also of limited use since they require fitting to the bulk data.

In 1993, Lew Yan Voon and Ram-Mohan [4] took a different approach. Drawing from $k \cdot p$ theory, it was realized that there had to be a relationship between the momentum matrix elements and the Hamiltonian matrix elements. Using the generalized Feynman-Hellmann theorem, we obtained an exact relationship between the two operators

$$\mathbf{p}(\mathbf{k}) = \frac{m_0}{\hbar}\nabla_{\mathbf{k}}H(\mathbf{k}). \tag{6}$$

The momentum matrix elements were then given by

$$
\begin{aligned}
\mathbf{p}_{nm}(\mathbf{k}) &= \frac{m_0}{\hbar} \langle n\mathbf{k}|\nabla_{\mathbf{k}}H(\mathbf{k})|m\mathbf{k}\rangle \\
&= \sum_{\alpha\beta bb'} C_{\alpha b}^*(n\mathbf{k})C_{\beta b'}(m\mathbf{k})\mathbf{p}_{\alpha\beta}^{bb'}(\mathbf{k}),
\end{aligned}
\tag{7}
$$

where

$$
\mathbf{p}_{\alpha\beta}^{bb'}(\mathbf{k}) = \frac{im_0}{\hbar} \sum_{\mathbf{R}_{bb'}} e^{i\mathbf{k}\cdot\mathbf{R}_{bb'}} \mathbf{R}_{bb'} E_{\alpha\beta}^{bb'}(\mathbf{R}_{bb'}).
\tag{8}
$$

Comparing with the procedures of Xu and Chang and Aspnes, one finds that

$$
\mathbf{p}_{\alpha\beta}^{bb'}(\mathbf{k}) = \sum_{\mathbf{R}_{bb'}} e^{i\mathbf{k}\cdot\mathbf{R}_{bb'}} \mathbf{p}_{\alpha\beta}^{bb'}(\mathbf{R}_{bb'}),
\tag{9}
$$

and

$$
\mathbf{p}_{\alpha\beta}^{bb'}(\mathbf{R}_{bb'}) = \frac{im_0}{\hbar} \mathbf{R}_{bb'} E_{\alpha\beta}^{bb'}(\mathbf{R}_{bb'}).
\tag{10}
$$

The key result of the exact expression is that $\mathbf{p}_{\alpha\beta}^{bb'}(\mathbf{R}_{bb'})$ has no contribution from intraatomic excitations, in contradiction to the results of Xu and Chang and Aspnes. Also, it shows that one can obtain the momentum matrix elements from the Hamiltonian matrix elements and that the exact wave functions are not required.

RESULTS

We now present numerical results based upon Eq. (6). We start by studying the zone-center momentum matrix elements since the latter are often quoted in the literature. In Table 2, we list the calculated momentum matrix elements using three popular TB parameter sets: VHD from Vogl, Hjalmarson, and Dow [6], SC from Schulman and Chang [7], and LS from Lu and Sham [8]. VHD and SC are nearest-neighbor sp^3s^* models, while LS is a second nearest-neighbor sp^3s model. It is clear that different TB parameter sets lead to different momentum matrix elements. One source of the differences is the different fitting of the energies. The other has to do with the finite basis set used [4, 10].

We next show some examples of dielectric functions and JDOS (Figs. 1 and 2) calculated using the TB method and the parameters of Vogl, Hjalmarson, and Dow. While there are many differences between the experimental and calculated spectra (particularly with respect to magnitude), the general shapes are reproduced. A common approximation in the literature is to substitute the JDOS for ϵ_2. Figs. 1 and 2 show that this is valid for Si but not for GaAs. The calculations of the dielectric functions for bulk zincblende semiconductors on a DEC Alpha workstation took a mere 15 mins for over 20000 special points.

For completeness, we show in Fig. 3 the kind of spectra obtained using different TB parameter sets for the second-harmonic generation coefficient [5]. The agreement between different parameter sets is less good than for linear optics.

CONCLUSION

We have reviewed and compared two distinct procedures for obtaining the momentum matrix elements within the TB formalism: a fitting procedure and the exact one. It has

379
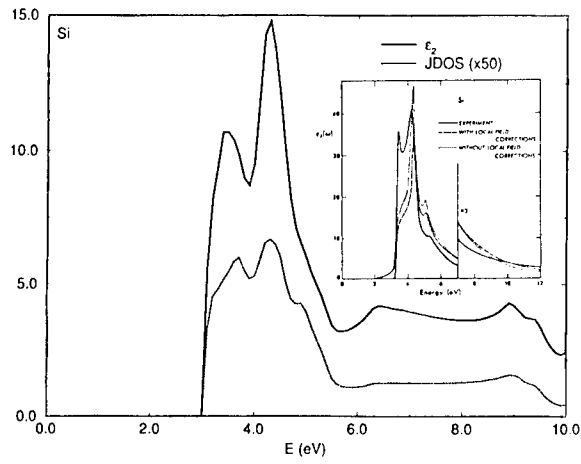
Figure 1: Imaginary part of dielectric function and JDOS of Si calculated using TB. Inset: experimental curve from Cohen and Chelikowsky [11].
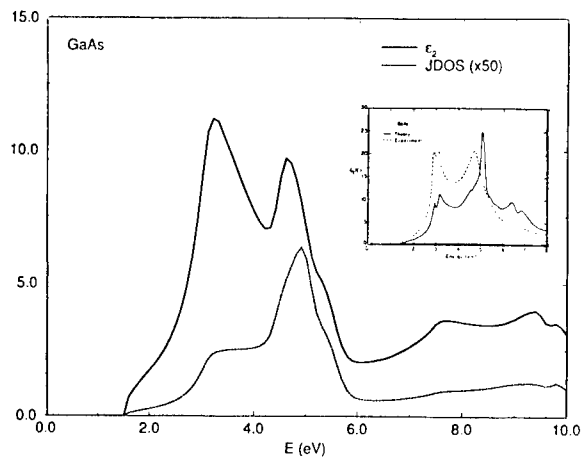
Figure 2: Imaginary part of dielectric function and JDOS of GaAs calculated using TB. Inset: experimental curve from Cohen and Chelikowsky [11].
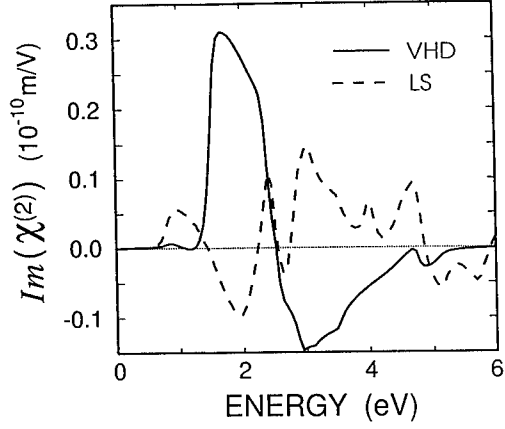
Figure 3: Dispersion of the imaginary part of the second-harmonic coefficient calculated using two different TB parameter sets.

Table 2: Interband matrix elements $E = 2P^2/m_0$ (in eV) and key band energies and effective masses for GaAs at the Γ point calculated from three different tight binding models. E_0 is the fundamental direct gap and E_1 is the Γ_{15c}-Γ_{1c} (Γ_{7c}-Γ_{6c}) gap. These are also compared with values given by Gorczyca *et al.* [9].

GaAs	VHD	SC	LS	Gorczyca
E_{P_0}	15.44	25.3	18.41	24.9,26.3,27.86
E_{P_1}	0.01	0.36	0.97	0.11,0.78,2.361
E_{Q_0}	13.52	11.94	12.53	13.4,13.7,15.56
E_0	1.55	1.43	1.52	1.66
E_1	3.16	3.02	3.20	3.21
Δ_0	—	0.343	—	0.338
Δ_1	—	0.221	—	0.176
m_n^Γ	0.119	0.067	0.067	0.066
$m_{hh}^\Gamma(\Delta)$	−0.409	−0.454	−0.511	—
$m_{lh}^\Gamma(\Delta)$	−0.089	−0.070	−0.055	—

been shown for the first time that the fittings so far carried out are unphysical. Examples of calculated linear and nonlinear susceptibilities are given.

REFERENCES

[1] J. C. Slater and G. F. Koster, Phys. Rev. **94**, 1498 (1954).

[2] Z. Xu, Solid State Commun. **76**, 1143 (1990).

[3] Y. C. Chang and D. E. Aspnes, Phys. Rev. B **41**, 12002 (1990).

[4] L. C. Lew Yan Voon and L. R. Ram-Mohan, Phys. Rev. B **47**, 15500 (1993).

[5] L. C. Lew Yan Voon and L. R. Ram-Mohan, Phys. Rev. B **50**, 14421 (1994).

[6] P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Solids **44**, 365 (1983).

[7] J. N. Schulman and Y. C. Chang, Phys. Rev. B **31**, 2056 (1985).

[8] Y.-T. Lu and L. J. Sham, Phys. Rev. B **40**, 5567 (1989).

[9] I. Gorczyca, P. Pfeffer, and W. Zawadzki, Semicond. Sci. Technol. **6**, 963 (1991).

[10] M. Graf and P. Vogl, Phys. Rev. B **51**, 4940 (1995).

[11] M. L. Cohen and J. R. Chelikowsky, Electronic Structure and Optical Properties of Semiconductors (Springer-Verlag, Berlin, 1988).

SUPERLATTICE CALCULATION IN AN EMPIRICAL $spds^*$ TIGHT-BINDING MODEL

R. SCHOLZ[(1)], J.-M. JANCU[(2)], and F. BASSANI[(2)]
(1) Institut für Physik, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, scholz@physik.tu-chemnitz.de
(2) Scuola Normale Superiore and Istituto Nazionale per la Fisica della Materia, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

ABSTRACT

We propose an empirical tight-binding method for tetrahedrally coordinated cubic materials and apply it to group IV and III-V semiconductors, extending existing calculations by the inclusion of all five d-orbitals per atom in the basis set. The symmetry character of the conduction states at the surface of the Brillouin zone is considerably improved compared to calculations in smaller bases, and the corresponding band positions can be obtained within the experimental uncertainties. Because the distance dependence of the tight-binding parameters is derived from deformation potentials, the model is particularly suited for an investigation of strained superlattices where the states at direct or pseudo-direct conduction band minima are composed of wavefunctions of all the minima at Γ, X, and L of the constituents. Investigations of GaAs/AlAs and short-period superlattices indicate a strong mixing between the conduction band valleys in the miniband structure, and the results are in better agreement with experiments than state-of-the-art empirical pseudopotential calculations.

INTRODUCTION

During the last decades, empirical tight-binding (TB) methods have been used for an approximate understanding of materials properties of various crystal symmetries [1-3]. A major problem for quantitative applications concerning cubic semiconductors has been the required extension of the atomic basis from the minimal size sp^3 [2,4] towards higher-lying atomic states like the next s-shell in an $sp^3 s^*$ basis [5], d-states of Γ_3-symmetry in $sp^3 d^2$ [6] and the recent parametrization including $sp^3 d^5 s^*$ states [7]. Group theoretical arguments demonstrate that the latter $sp^3 d^5 s^*$ basis is needed for a good approximation to numerical completeness, and various deficiencies of the smaller TB models have been resolved, e.g. the decomposition of valence and conduction wavefunctions into different atomic symmetries of cation and anion [7].

A crucial step in the parametrization of our $sp^3 d^5 s^*$ nearest-neighbour TB model was the investigation of the free-electron band structure along the lines indicated in older empirical TB calculations. This allowed an analytic derivation of universal parameters for zincblende- and diamond-like semiconductors. In particular, the on-site energies of the d-states and all their interaction parameters were deduced from the higher free-electron conduction bands [7]. As these bands remain free-electron like in cubic semiconductor materials, the corresponding TB parameters are typical for all materials of this symmetry class. Because the d-d-interaction parameters interfere constructively at some points of the Brillouin zone, the corresponding splittings are by far too large to be obtained in low order perturbation theory.

The lowest bonding d-state at the high symmetry points of the Brillouin zone occurs at X, leading to a dominating d-contribution to the X_{6c} wavefunction. The small transverse conduction mass at X is related to p-d-interactions and cannot be reproduced in nearest-neighbour TB models excluding d-states. Owing to the important admixture of $\Gamma_4(T_{2g})$-symmetric d-states to the highest valence band at Γ, the second conduction band at Γ and to the X_{6c} valley, the use of a $sp^3 d^2$ basis including only the $\Gamma_3(E_{1g})$-symmetric d-states [6]

383

leads to severe shortcomings, resulting e.g. in poor estimates for dipole matrix elements and masses. The s^*-state, on the contrary, is not really needed for good agreement with the experimental band structure. It is useful, however, in order to conserve the correspondence of the lower Γ_1-symmetric s-state with the free atom, and therefore the transferability of the on-site energies for the two valence states between different compounds, like e.g. E_p^a and E_s^a for As in the two materials GaAs and AlAs investigated here. Furthermore, the anion-cation character of Γ_{6c} would not correspond to first-principles results in the smaller sp^3d^5 basis, leading to severe shortcomings in heterostructures where the localization of the wavefunctions on well-defined atomic planes is of crucial importance [8,9].

FIG. 1:
Comparison of GaAs and AlAs bulk band structures, in eV. The calculated (experimental, Ref. [10]) energies for the conduction minima, measured from the valence band maximum, are:
GaAs:
$E(\Gamma_{6c}) = 1.519$ eV (1.519 eV),
$E(L_{6c}) = 1.837$ eV (1.85 eV),
$E(X_{6c}) = 1.989$ eV (1.98 eV),
$E(X_{7c}) = 2.328$ eV (2.35 eV).
AlAs:
$E(\Gamma_{6c}) = 3.130$ eV (3.13 eV),
$E(L_{6c}) = 2.581$ eV (2.54 eV),
$E(X_{6c}) = 2.223$ eV (2.229 eV),
$E(X_{7c}) = 2.584$ eV (2.579 eV).
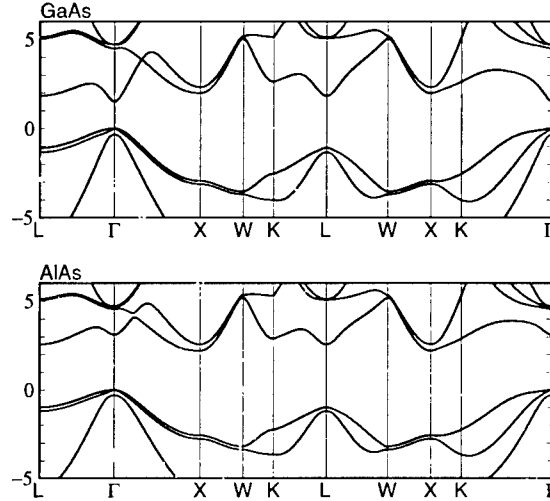
The band structures resulting in our $sp^3d^5s^*$ TB model for GaAs and AlAs are shown in Fig. 1. The conduction band energies at the minima Γ_{6c}, L_{6c}, X_{6c} and X_{7c} as well as the corresponding masses $m(\Gamma_{6c})$ and $m_t(X_{6c})$ are reproduced within the experimental uncertainties. The orderings $E(\Gamma_{6c}) < E(L_{6c}) < E(X_{6c})$ for GaAs and the reversed ordering for AlAs can be related to the size of the cation core: The d-dominated wavefunction at X_{6c} is pushed up by the presence of d-states in the Ga-core, an effect not present in AlAs. Using the valence band maximum as the conventional energy zero, this effect is not directly reflected in the TB parameters because the energies of the empty d-states and their interactions remain quite close in both materials, resulting in similar X_{6c}-energies. The main changes between the TB parameters of the two materials are a shift of the on-site energy E_s^c of about 1.4 eV upwards between GaAs and AlAs, and a significant change in the interaction $s_a^* s_c \sigma$, leading to a much higher Γ_{6c}-state in AlAs [7].

ELECTRONIC BAND STRUCTURE OF $(GaAs)_n(AlAs)_n$ SUPERLATTICES

We demonstrate the quality of the present TB model by calculating the electronic properties of short period $(GaAs)_n/(AlAs)_n$ [001] superlattices and comparing the results with experimental data and pseudopotential (PP) calculations. The lattice constant orthogonal to the growth direction is fixed to the GaAs substrate, resulting in a compression of the AlAs lattice in these plains and correspondingly a dilatation along the growth direction. Calculating the strain in the AlAs layers from elasticity theory, the modifications of the

TB parameters are analogous to bulk material under tetragonal strain. As the interface between GaAs and AlAs is always an As layer, the bonds of the As atoms to Ga and Al can be parametrized as in the corresponding bulks, except for the strain corrections. The on-site energies of all atoms not on the interface are shifted between the two materials in accordance with the valence band offset. On the other hand, the on-site energies of the atomic-like orbitals on the As interface layers take the average values between the two bulks, shifted therefore by half the valence band offset compared to As-sites in the central part of the GaAs and AlAs layers. As the valence band offset cannot be derived from our empirical TB calculation, it is introduced as an additional model parameter characterizing the interface, chosen as $E_v = 0.55$eV in accordance with experimental results [11].

FIG. 2:
Comparison of experimental and tight-binding gap energies in $(GaAs)_n(AlAs)_n$ superlattices. States deriving from Γ: TB (solid), experimental (open $\diamond$: low T [12], filled $\diamond$: RT [13], open $\Box$: compilation of different temperatures [14]); X_z: TB (long dashes), experimental (open $\circ$: low T [12], filled $\circ$: RT [13]); $X_{x,y}$: TB (dash-dotted), experimental ($\times$, low T [12]); L_{6c}: TB (short dashed).
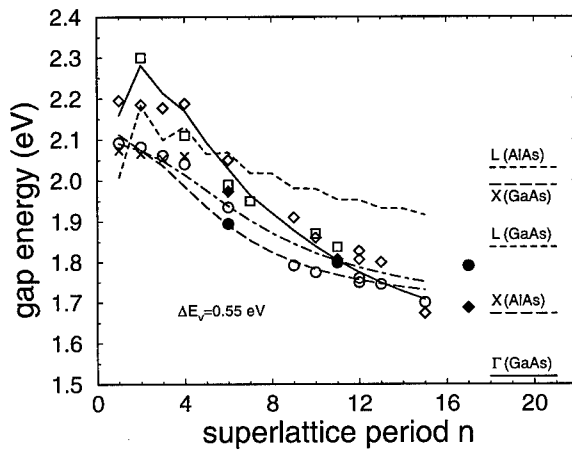
Fig. 2 shows the calculated direct and indirect gap energies together with experimental data. The TB gaps correspond to free electron-hole recombination, neglecting excitonic effects. We found the $n = 1$ superlattice is characterized by a L-derived conduction band minimum (CBM) mainly localized on the Ga sublattice ($s(Ga) + p(Ga) + d(Ga) \sim 40\%$), in agreement with first-principle calculations [15]. Both in empirical and *ab initio* pseudo-potential (PP) investigations [8,15] and in the present tight-binding approach, this energy level is found to oscillate strongly with the superlattice period n. Quantum-confinement effects at the L-point are most pronounced within the empirical PP method owing to a higher electron potential barrier: These calculations overestimate the difference of the L_{6c} energies in bulk AlAs and GaAs by 0.33 eV [8]. For $n = 2$, the TB CBM derives from the bulk AlAs $X_{x,y}$- states because tetragonal strain shifts the X_z-level above $X_{x,y}$ by about 20 meV for $n = 1$ and 2, in agreement with experiment [12]. As the corresponding PP calculations neglect strain, this is not observed [8]. Due to the large longitudinal mass $m_l(X)$, the quantum confinement energy of the X_z states is smaller than for the $X_{x,y}$ states confined according to the much smaller transverse mass $m_t(X)$. This confinement effect induces a $X_{x,y} \to X_z$ crossover for the CBM between $n = 2$ and 3 in our calculations, in good agreement with experiment, where this crossover occurs between $n = 3$ and 4 [12]. For $3 \leq n \leq 13$, the TB CBM derives from the bulk AlAs X_z-states folded to the center of the tetragonal Brillouin zone. Since the valence band maximum (VBM) is a Γ-like state localized in the GaAs well, the interband transition is then pseudodirect in momentum space and indirect in real space (type-II). The quantum size effects associated with the period n induce a type-II $\to$ type-I transition between $n = 13$ and 14 monolayers, and

for $n \geq 14$ the CBM is a Γ-like state localized mainly in the GaAs well. This calculated crossover thickness is in good agreement with the experimental results for the crossover thickness $n_c \approx 11$ at room temperature [13] and $n_c \approx 14$ at low temperature [12]. The overall agreement between theory and experiment is excellent and clearly demonstrates the transferability of the TB parameters to quantum structures with energy offsets at the interfaces. Interestingly, the best empirical pseudopotential approaches use As pseudo-potentials depending on the number of neighbouring Ga and Al sites [8], in qualitative correspondence with our TB parametrization of the interface. However, the advantage of the TB model is that we can better distinguish the bonds towards the Ga and Al sites: They are parametrized as in the corresponding (strained) bulks.

Both type-II and type-I transitions refer to the same heavy hole VBM, so that the type-II $\rightarrow$ type-I crossover at $n_c \approx 11$ is a pure effect of the folding of the conduction minima in the superlattice. Therefore, it is most instructive to compare the conduction band energies calculated in different approaches. In order to guarantee comparability between our TB calculation and pseudopotential results, we use now the same valence band offset of $\Delta E_v = 0.50$ eV [8,9]. This reduces the type-I type-II crossover thickness in our TB calculation to $10 < n_c < 11$, compare Fig. 3.

FIG. 3:
Comparison of calculated energies for the conduction minima in $(GaAs)_n(AlAs)_n$ superlattices. States deriving from bulk Γ_{6c} states (lowest 3 lines for large n): TB (solid), PP (dotted) [9], $k \cdot p$ (short dashes) [9]; X_z: TB (long dashes), PP (dotted); $X_{x,y}$: TB (dash-dotted), PP(dotted); L_{6c} (highest 2 lines for large n): TB (short dashed), PP (dotted). The type-II (X_z) type-I (Γ) crossing thicknesses in the PP and TB calculations are indicated by circles, and the bulk and average energies of the conduction minima as horizontal bars.
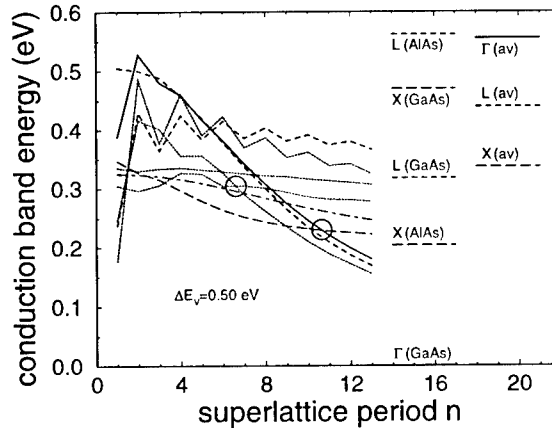
For large layer thicknesses, the superlattice minima approach the lower of the bulk valleys (GaAs for Γ and L, AlAs for X), with a confinement energy proportional $1/(n^2 m^*)$. Interestingly, the superlattice minima for layer thicknesses $n = 2 - 3$ are quite close to the average of the corresponding bulk values. The confinement in the well material leads to a small increase of the wavefunction within the well, and correspondingly to a quite small reduction of this average energy. The too low crossover thickness $6 < n_c < 7$ in empirical PP calculations [8,9] can be related to a too small confinement energy, scaling inversely with the GaAs bulk mass of $m_c(\Gamma) = 0.099$ instead of the smaller experimental value of $m_c(\Gamma) = 0.067$ [8]. We suppose an empirical PP parametrization resulting in an improved value for the bulk-GaAs conduction band mass would be in much better agreement with the experimental crossover thickness $n_c \geq 11$ and our TB value of $10 < n_c < 11$.

$k \cdot p$ perturbation theory in the envelope approximation [9] shows a monotonous decrease of the Γ_{6c} confinement energy with increasing layer thickness n. It does neither reproduce the small even-odd oscillation nor the correct bands deriving from X and L, which are about 20 eV too high. Therefore, it is only applicable for type-I quantum structures with sufficient layer thickness, compare the detailed discussion in Ref. [9].

INTER-SUBBAND TRANSITIONS

Recently, X-Γ intersubband resonances [16] have been observed in GaAs/AlAs [001] superlattices under transverse magnetic (TM) polarization with the electric light field partially along the [001] growth direction. In dipole approximation, such transitions related to z-polarized light are forbidden in zincblende-type crystals, but in superlattices, they become weakly allowed due to intervalley mixing effects at the interfaces. In order to gain some insight into the underlying mechanism, we have calculated the momentum matrix elements of $(GaAs)_n/(AlAs)_n$ [001] superlattices for z-polarized light at the center of the tetragonal Brillouin zone.

FIG. 4:
Energy gap $E(\Gamma) - E(X_z)$ (long dashes), dipole matrix element P (a.u.) (short dashes with o), and optical transition strength (a.u.) (solid). The valence band offset used in this calculation is $\Delta E_v = 0.55$ eV.
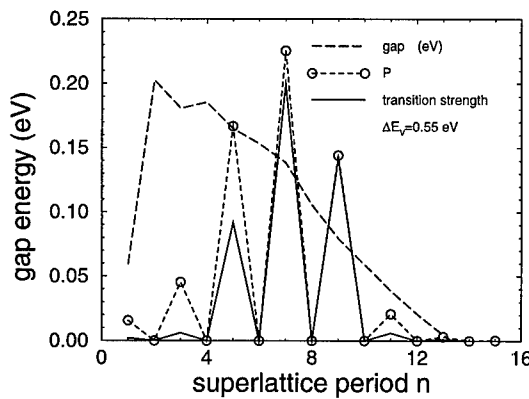
Fig. 4 shows the calculated oscillator strength between X_z-like and Γ-like states. Such optical transitions remain dipole-forbidden for even n, suggesting that the Γ-X mixing strongly depends on the parity of n. These features can be understood qualitativeley by considering the superlattice states in terms of bulk states and zone folding. Because of zone folding, conduction states at the superlattice zone center can evolve either from a bulk Γ state, $\overline{\Gamma}(\Gamma_{1c})$, or from non-bulk states $\overline{\Gamma}(X_{1c})$ and $\overline{\Gamma}(X_{3c})$, the latter two with opposite parity with respect to the central atomic layer of each material. For even n, the central atomic plane in the AlAs barrier is occupied by As, and since the X_{1c} states are invariant under mirror reflections $S_4(As)$ with the origin on the anion site, the folded states $\overline{\Gamma}(\Gamma_{1c})$ and $\overline{\Gamma}(X_{1c})$ have the same symmetry, $\overline{\Gamma_{1c}}$. This ensures that the momentum matrix element $\langle \overline{\Gamma_{1c}}(\Gamma_{1c})|\mathbf{p}|\overline{\Gamma_{1c}}(X_{1c})\rangle$ is zero since the momentum operator transforms like Γ_4. For odd n, X_{1c} is an anti-symmetric state with respect to the central Al layer, so that it folds into the superlattice state $\overline{\Gamma_{4c}}(X_{1c})$. Moreover, for odd n the states $\overline{\Gamma}(\Gamma_{1c})$ and $\overline{\Gamma}(X_{3c})$ have the same symmetry $\overline{\Gamma_{1c}}$ and retain therefore both the character of bulk Γ_{1c} and X_{3c}-states. The momentum matrix element $\langle \overline{\Gamma_{1c}}(\Gamma_{1c})|\mathbf{p}|\overline{\Gamma_{4c}}(X_{1c})\rangle$ is non-zero in this case, and the corresponding intersubband transition becomes dipole-allowed. Its intensity depends on the strength of the Γ_{1c}-X_{3c} mixing and is particularly enhanced for ultrathin layers with significant quantum confinement effects, i.e. true two-dimensional states. Note that this dipole-allowed optical transition has a counterpart in the bulk since $\langle X_{3c}|\mathbf{p}|X_{1c}\rangle$ is non-zero (**p** transforms like $X_5 + X_3$ at X).

For small and large n, the wavefunctions remain close to the corresponding bulks and behave like three-dimensional states. The mixing between Γ_1- and X_3-derived superlattice states is not pronounced, so that the dipole matrix elements in Fig. 4 remain small. Only for intermediate thicknesses n, the wavefunctions show a two-dimensional character with larger admixture of Γ_1 and X_3, resulting in larger dipole matrix elements. For small n,

the gap is large, so that in addition to the small Γ_1-X_3 mixing, the transition strength is suppressed by the large energy denominator.

Recent experiments [16] have suggested that TM absorption in $(GaAs)_6/(AlAs)_8$ [001] superlattices are due to X-Γ intersubband transitions. However, our calculations clearly evidence that these transitions are dipole-forbidden, as explained with the symmetry analysis carried out before. Accordingly, a new interpretation of the experiment is necessary, including band-band transitions away from the high-symmetry points. A calculation of the dielectric function and the inter-subband absorption coefficient based based on our TB model is in progress, compare [17] for interband absorption in heterostructures. Possibly, the p-n-junction around the superlattice in the experimental set-up leads to a significant electric field, resulting in a weakening of the selection rules derived from group theory.

CONCLUSION

The $spds^*$ basis extension in our empirical TB model for zincblende materials has transformed this method into a highly predictive tool, both for bulk materials and for quantum structures. The comparison of our superlattice bands with experimental gaps suggests that a valence band offset of $\Delta E_v = 0.55$ eV [11] is quite realistic for GaAs/AlAs interfaces. The method presented in this contribution is well suited for computations where the diagonalization of the Hamiltonian is the most time-consuming step, like e.g. investigations of the self-consistent field in heterostructure devices under forward bias [17].

REFERENCES

1. J.C. Slater and G.F. Koster, Phys. Rev. 94, 1498 (1954).
2. D.J. Chadi and M.L. Cohen, phys. stat. sol. (b) 68, 405 (1975).
3. J.A. Majewski and P. Vogl, Phys. Rev. B 35, 9666 (1987).
4. W.A. Harrison, Electronic Structure and Properties of Solids, (Freeman, San Francisco 1980).
5. P. Vogl, H.P. Hjalmarson, and J.D. Dow, J. Phys. Chem. Solids 44, 365 (1983).
6. Y.C. Chang and D.E. Aspnes, Phys. Rev. B 41, 12002 (1990).
7. J.-M. Jancu, R. Scholz, F. Beltram and F. Bassani, Phys. Rev. B, in press.
8. K.A. Mäder and A. Zunger, Phys. Rev. B 50, 17393 (1994).
9. D.M. Wood and A. Zunger, Phys. Rev. B 53, 7949 (1996).
10. Semiconductors: Group IV Elements and III-V Compounds, ed. by O. Madelung (Springer, Berlin 1991)
11. J. Batey and S.L. Wright, Surf. Sci., 174, 320 (1986).
12. M. Holtz, R. Cingolani, K. Reimann, R. Muralidharan, K. Syassen, and K. Ploog, Phys. Rev. B 41, 3641 (1990); W. Ge, W.D. Schmidt, M.D. Sturge, L.N. Pfeiffer, and K.W. West, J. Lumin. 59, 163 (1994); M. Nakayama, K. Imazawa, K. Suyama, I. Tanaka, and H. Nishimura, Phys. Rev. B 49, 13564 (1994).
13. G. Li, D. Jiang, H. Han, Z. Wang, and K. Ploog, Phys. Rev. B 40, 10430 (1989).
14. R. Cingolani, L. Baldassare, M. Ferrara, M. Lugarà, and K. Ploog, Phys. Rev. B 40, 6101 (1989).
15. S.-H. Wei and A. Zunger, J. Appl. Phys. 63, 5794 (1988).
16. A. Fenigstein, E. Finkman, G. Bahir, and S.E. Schachman, Appl. Phys. Lett. 69, 1759 (1996).
17. A. Di Carlo, this volume

SELF-CONSISTENT TIGHT-BINDING METHODS APPLIED TO SEMICONDUCTOR NANOSTRUCTURES

ALDO DI CARLO
INFM-Dipartimento di Ingegneria Elettronica, Università di Roma "Tor Vergata",
00133 Roma, Italy. dicarlo@eln.utovrm.it

ABSTRACT

A self-consistent tight-binding approach applied to semiconductor nanostructure is presented. This allows us to describe electronic and optical properties of nanostructured devices beyond the usual envelope function approximation. Example of applications are given for High Electron Mobility Transistors (HEMTs) and non-linear optical devices.

INTRODUCTION

Modern microelectronics and optoelectronics is heavily based upon semiconductor nanostructures, where the dimension of the "active" region is in the nanometric range. Typical examples are the High Electron Mobility Transistor (HEMT), the semiconductor LASER, the semiconductor optical amplifiers (SOA). Basic element of all these structures is the heterojunction between different semiconductor materials.

The theoretical study of electronic and optical properties of such devices has been undertaken by different methods. These range from *ab-initio* approaches [1], which are very precise but require a large computational effort and, consequently, are limited only to very small nanostructures, to approximate but easy-to-handle and fast methods such as those based on the k.p expansion in the envelope function approximation (EFA) [2]. In its simplest form, the EFA leads to the evaluation of the energy levels of nanosystems by simply solving a one-electron Schrödinger equation where each semiconductor is described in terms of effective masses and band edges. [3] Despite its easy applicability, EFA is limited by several factors. First of all, it is not capable to describe the nanosystem in the whole Brillouin zone. Second, the use of the same periodic part of the Bloch function for all the semiconductors forming the heterojunction makes EFA unapplicable to very thin structures. Moreover, the level of description needed to account for and to predict the properties of last generation micro and optoelectronic devices cannot be reached within the context of simple EFA methods.

The empirical tight-binding method (TB) [4] has been shown to be a valid alternative to EFA, since it improves the physical content in the description of the nanostructure with respect to EFA without requiring a much higher computational effort. In particular, it allows us to treat indirect-gap semiconductors, heterostructures formed by indirect and direct materials, and to describe very thin layers [5, 6, 7, 8].

TB has been mainly used in the calculation of the electronic properties of nanostructures without taking into account self-consistent charge redistribution, which is an important requirement when we deal with real systems. However, very recently, we have shown [9] that a self-consistent tight-binding procedure can be defined.

In this paper we will describe the self-consistent TB model and its application to the study of the electronic and optical properties of realistic nanostructures.

389

THEORY

In this section we discuss the self-consistent tight-binding model for a system where the translational symmetry is broken in one direction, for example the growth axis (z). The wave function $|E, \mathbf{k}_\parallel\rangle$ can be written as linear combination of planar Bloch sums, $|\alpha, m, \mathbf{k}_\parallel\rangle$ [7, 10]

$$|E, \mathbf{k}_\parallel\rangle = \sum_{\alpha, m} C_{\alpha,m}(E, \mathbf{k}_\parallel)|\alpha, m, \mathbf{k}_\parallel\rangle \qquad (1)$$

with

$$|\alpha, m, \mathbf{k}_\parallel\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}^m_{\alpha\parallel}} e^{i\mathbf{k}_\parallel \cdot \mathbf{R}^m_{\alpha\parallel}} |\alpha, \mathbf{R}_\alpha\rangle , \qquad (2)$$

where $|\alpha, \mathbf{R}_\alpha\rangle$ is a localized orbital, $\mathbf{k}_\parallel$ is the in-plane wave vector and N is the number of unit cells in the atomic plane. The subindex α refers both to the basis atom index and to the atomic orbital index. The lattice vector, $\mathbf{R}_\alpha = \mathbf{R} + \mathbf{v}_\alpha$, (where $\mathbf{v}_\alpha$ is the basis atom displacement), can be written as $\mathbf{R}_\alpha = m\,\mathbf{d} + \mathbf{R}^m_{\alpha\parallel}$ where m is an integer, $\mathbf{d}$ a vector parallel to the growth direction with module equal to the distance between two atomic planes and $\mathbf{R}^m_{\alpha\parallel}$ is a vector on the m-th atomic plane. For a given $\mathbf{k}_\parallel$, the eigenstates E are calculated by solving the secular equation $(H + V_H)|E, \mathbf{k}_\parallel\rangle = E|E, \mathbf{k}_\parallel\rangle$ where H is the system tight-binding hamiltonian and V_H is the Hartree potential. The influence of the electronic charge rearrangement can be included at a Hartree level by solving the Poisson equation, $d^2V_H/dz^2 = -\rho(z)/\varepsilon$, where ε is the static dielectric constant. The charge density in the m-th plane $\rho(m)$ is defined by:

$$\rho(m) = -\frac{e}{(2\pi)^2} \int_{BZ_\parallel} d\mathbf{k}_\parallel \sum_{n,\alpha} \left|C_{n,\alpha}(E_n, \mathbf{k}_\parallel)\right|^2 \tilde{f}(E_n, E_F) \qquad (3)$$

where e is the electron charge and n labels the energy levels for a given $\mathbf{k}_\parallel$. The function $\tilde{f}(E_n, E_F)$ is defined as follows

$$\tilde{f}(E_n, E_F) = \begin{cases} f(E_n, E_F) & \text{for the conduction states} \\ 1 - f(E_n, E_F) & \text{for the valence states} \end{cases} , \qquad (4)$$

where $f(E_n, E_F)$ is the Fermi distribution function with a given Fermi level E_F. $\tilde{f}(E_n, E_F)$ is a well behaved function which is different from zero only in the proximity of the valence and conduction band edges. In order to evaluate the carrier density (Eq. 3), a full $\mathbf{k}_\parallel$ integration is performed in the 2D Brillouin zone ($BZ_\parallel$) by using the special k-points technique in the irreducible wedge [11]. The convergence of this integration has been obtained by using 5 special points with $|\mathbf{k}_\parallel| \leq 0.06\ 2\pi/a$ for direct band gap material. To achieve self-consistency in indirect band gap material we use 8 points with $|\mathbf{k}_\parallel| \leq 0.2\ 2\pi/a$.

Poisson and Schrödinger equations in the TB representation are iteratively solved until convergence is reached. Open-chain (infinite well) boundary conditions are used for Schrödinger's equation. In order to avoid and influence on the calculated electronic levels, boundaries are chosen far away from the nanostructure active region. This may not be satisfactory for high energy states such as those above barrier. Indeed, a better choice for the boundary condition is provided by applying the scattering theory as explained in Ref. [10]. However, for all the situations discussed here the open chain condition represents a valid choice. Moreover, the use open chain boundary condition induces the hamiltonian matrix to have a "band" form. This implies that very efficient diagonalization methods,

suited for this matrix form, can be applied. We have here introduced a hybrid method to diagonalize the tight–binding hamiltonian which uses a standard (LAPACK [12]) routine to calculate eigenvalues and an inverse iteration scheme [13] to calculate eigenvectors. The advantage of this procedure relies on the fact that only few eigenvectors are needed, namely those closed to the energy band gap. To speed-up the self-consistent TB calculation, we take as initial starting potential the one obtained self-consistently in the effective mass approximation.

When optical properties are of interest, one can make use of the Kubo formula to define the susceptibility tensor which is related to the current–current response function of the electromagnetic perturbation. This can be easily calculated within the tight-binding scheme *without* introducing new fitting parameters [14]. If we consider a linear polarization of the light along the i-th axis, the absorption coefficient can be written as [14, 9]

$$\alpha(\omega) = \frac{4\pi^2}{nc\omega S} \sum_{\substack{\alpha,m \\ \alpha',m'}} \sum_{E,E',\mathbf{k}_{||}} [f(E) - f(E')] \ \delta(\hbar\omega + E - E') |\langle E, \mathbf{k}_{||}| \sum_{\alpha,m} J_i(\alpha,m)|E',\mathbf{k}_{||}\rangle|^2 \quad (5)$$

Here, S is the transverse area of the primitive cell, n is the refractive index and c the speed of light. The matrix elements of the current operator can be expressed as:

$$\langle E, \mathbf{k}_{||}|\mathbf{J}(\alpha,m)|E',\mathbf{k}_{||}\rangle = \frac{S}{N} \sum_{\substack{\alpha',m' \\ \alpha'',m''}} C^*_{\alpha',m'}(E,\mathbf{k}_{||}) C_{\alpha'',m''}(E',\mathbf{k}_{||})$$

$$\times \sum_{\mathbf{R}^m_{\alpha||},\mathbf{R}^{m'}_{\alpha'||},\mathbf{R}^{m''}_{\alpha''||}} e^{i\mathbf{k}_{||}(\mathbf{R}^{m''}_{\alpha''} - \mathbf{R}^{m'}_{\alpha'})} \langle \alpha',\mathbf{R}_{\alpha'}|\mathbf{j}(\alpha,\mathbf{R}_\alpha)|\alpha'',\mathbf{R}_{\alpha''}\rangle, \quad (6)$$

where

$$\langle \alpha',\mathbf{R}_{\alpha'}|\mathbf{j}(\alpha,\mathbf{R}_\alpha)|\alpha'',\mathbf{R}_{\alpha''}\rangle = \frac{ie}{2\hbar dS} t_{\alpha',\alpha''}(\mathbf{R}_{\alpha'} - \mathbf{R}_{\alpha''}) \left[\mathbf{R}_{\alpha'} - \mathbf{R}_{\alpha''}\right] \left\{\delta_{\alpha'\mathbf{R}_{\alpha'},\alpha\mathbf{R}_\alpha} + \delta_{\alpha''\mathbf{R}_{\alpha''},\alpha\mathbf{R}_\alpha}\right\}.$$

Here $t_{\alpha',\alpha''}(\mathbf{R}_{\alpha'} - \mathbf{R}_{\alpha''}) = \langle \alpha',\mathbf{R}_{\alpha'}|H|\alpha'',\mathbf{R}_{\alpha''}\rangle$ is the tight-binding hopping matrix element.

The $\mathbf{k}_{||}$ integration needed to calculate the luminescence or absorption coefficient (Eq. 5) is performed in the irreducible wedge of the 2D Brillouin zone. Since, we are only interested to calculate these optical properties close to the energy gap, we limited the integration only to the region $|\mathbf{k}_{||}| < 0.1 \ 2\pi/a$.

In order to be able to change the carrier distribution function without recalculating all the eigenvalues/eigenvectors of the hamiltonian, we first calculate (and store) the energy levels and the squared optical matrix elements for each $\mathbf{k}_{||}$. We then evaluate the luminescence or the absorption coefficient by performing the sums in Eq. 5. To reduce the numerical fluctuation induced by the finite number of $\mathbf{k}_{||}$ points considered ($\sim$1600), we sum over a much finer $\mathbf{k}_{||}$ grid ($\sim 10^5$ points). Energy levels and squared matrix elements at these new $\mathbf{k}_{||}$ points are obtained by using a bilinear interpolation of the calculated quantities. This is allowed since variation of both energy levels and squared matrix elements in the irreducible wedge are quite smooth.

The strain is included in the TB model by scaling the hopping matrix elements by using the modified Harrison scaling law [15]. Tight-binding parameters are adapted from Ref. [7].

RESULTS

We consider a typical pseudomorphic (PM) HEMT device which consists of a 35 nm n-doped ($n=2\cdot 10^{18}\text{cm}^{-3}$) $Al_{0.2}Ga_{0.8}As$ cap layer, 2 nm undoped $Al_{0.2}Ga_{0.8}As$ spacer, 12 nm $In_{0.2}Ga_{0.8}As$ 2D channel, 4 nm undoped $Al_{0.2}Ga_{0.8}As$ spacer, 8 nm n-doped ($n=2\cdot 10^{18}\text{cm}^{-3}$) $Al_{0.2}Ga_{0.8}As$, 400 nm undoped $Al_{0.2}Ga_{0.8}As$ and finally a GaAs substrate. In this calculation the background doping charge is simply added to the free charge as given by Eq. 3. The self-consistent band edge profile for a $V_{gate} = 0$ is shown in Fig. 1 for both conduction and valence band. Here we have assumed a 0.8 eV Schottky barrier between AlGaAs and gate metal contact. Since there is not charge transport across the heterojunction (i.e. along the growth direction z) the Fermi level in the z direction is constant and is fixed from the Fermi level in the substrate and in the metal. Strain in the InGaAs region splits the heavy hole HH and the light hole LH bands. Here the labels HH and LH refer to the character of the valence band in the growth direction.

Under bias conditions close to the breakdown of the device, there is generation of holes due to impact ionization processes [16] which radiatively recombine with electrons around the gate region of the channel. The calculated photoluminescence spectra is shown in Fig. 2. Since the HH quantized levels are now above the LH states, the luminescence transitions occur between conduction levels and the heavy hole levels. The main peak corresponds to the $C_1 \rightarrow HH_1$ transition, while the other two peaks are related to the $C_2 \rightarrow HH_1$ and $C_2 \rightarrow HH_2$ transitions. The $C_2 \rightarrow HH_1$ transition, forbidden in a flat infinite barrier quantum well [2], is in this case allowed due to the symmetry breaking (consequence of the band bending) in the channel [9]. The emitted light is mainly polarized in the in-plane direction since the levels have essentially a heavy-hole character.

In the previous case, all semiconductors forming the device have a direct gap. The strength of this self-consistent TB approach consist in the possibility to treat at the same time systems formed by indirect gap semiconductors. In order to compare the TB with EFA in these situations we consider an AlGaAs/GaAs lattice matched HEMT. The Al concentrations of the cap layer is chosen in order to have $E_\Gamma = E_X$. It becomes evident (see Fig. 3) that EFA completely fails in describing the charge redistribution between the 2D channel and the parasitic channel which forms due to the large X-valley contribution to the charge density.
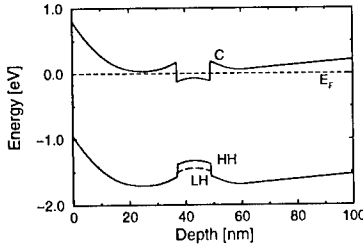
Fig.1 Self-consistent band edge profile of the PM HEMT. The dashed line represent the position of the Fermi level.
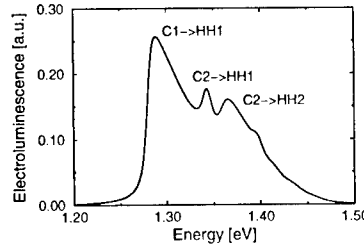
Fig.2 HEMT electroluminescence arising from the recombination of electrons with holes created by the impact ionization process. A Lorentzian broadening of 4 meV has been considered.

392

Since the self-consistence is achieved regardless of the type of the energy gap, the self-consistent TB approach can be also used in Si/SiGe devices where conventional device simulators are unable to describe the electronic properties [17]

In the HEMT examples given above, the position of the Fermi level was given by alignment condition between substrate and gate contact. However, the developed approach allows to self-consistently calculate the Fermi level as a function of the excess electrons(holes) in the conduction(valence). Systems where this approach is needed are, for example, those where the charge created by an optical excitation induces an electrostatic potential which modifies the bending profile of the structure and, in turn, its optical properties. Such devices are very useful for their optical non-linearities and are used for example as optical switches, frequency converters and optical amplifiers [18].

The investigated nanostructure consists on a $InP/In_{0.53}Ga_{0.47}As$ quantum well. In the middle of the InGaAs well a strained GaAs layer is inserted which behave as a barrier for electrons and a well for holes.
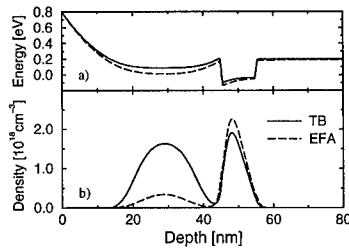
Fig.3 (a) Self consistent band profile of a $Al_{0.45}Ga_{0.55}As/GaAs/Al_{0.3}Ga_{0.7}As$ HEMT structure. (b)Self consistent charge density of the HEMT
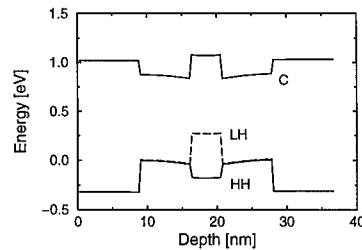
Fig.4 Self consistent band profile of the optical switch structure under high intensity laser pumping.

The self consistent band edge profile is shown in Fig. 4 for a high pumping laser intensity. The laser intensity fixes the injected electron charge density in conduction (equal to the hole density in valence) which is used as input parameter in our calculation. A more rigourous account for the injection mechanism can be found elsewhere [18]. From the knowledge of the pumping density and the energy levels (which is obtained via TB) the position of the quasi Fermi levels is calculated. By increasing further the laser intensities, electrons are promoted to the conduction band . However, electrons move in the InGaAs region while holes will be confined in the GaAs layer, as depicted if Fig. 5. This charge separation induces a net dipole moment which is responsible for band bending. The position of the quasi Fermi level with respect to the scale shown in Fig. 4 is $E_{Fc} = 0.9279$ eV for electrons and $E_{Fc} = 0.0722$ eV. The absorption coefficient of the structure with and without laser pumping is shown in Fig. 6. The effect of the laser is to create population inversion, thus a negative absorption (gain) is achieved. Beside this, we notice how the structure of the absorption is changed due to the redistribution of the charges. Indeed, the onset of the absorption is shifted by more than 30 meV with respect to the situation without laser pumping. Moreover, the different number of peaks between the two situations reflects the changing of levels position and oscillator strength.

CONCLUSIONS

In conclusion, we have shown that a self-consistent tight-binding approach can be used to evaluate the electronic structure and optical properties of semiconductor nanostructures. This represents a further step with respect to the envelope function model towards an *ab initio* calculation of such properties.
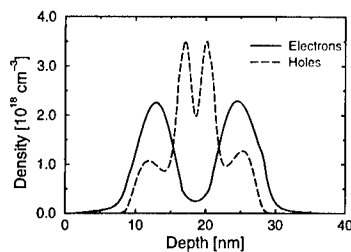
Fig.5 Self-consistent electron and holes charge density of the optical switch.
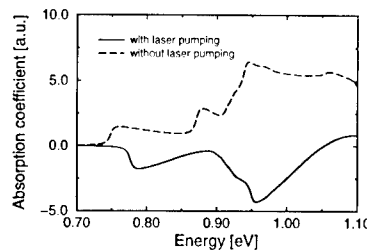
Fig.6 Absorption coefficient for the optical switch with (see Fig.4) and without (flat bands) laser pumping

REFERENCES

1. D. M. Wood, S.-H. Wei, and Alex Zunger, Phys. Rev. **B37**, 1342 (1988); C. H. Park and K. J. Chang, Phys. Rev. **B47**, 12709 (1993).
2. G. Bastard, Wave mechanics applied to semiconductor heterostructures, Les Edition de Physique, Les Ulis Cedex, 1988.
3. H. Morkoç, H. Unlu, G. Ji, Principles and Technology of MODFETS, John Wiley & Sons, Chichester, 1991.
4. J. C. Slater, G. F. Koster, Phys. Rev. **94**, 1498 (1954); D. W. Bullet, Solid State Physics, **35**, 129 (1980); J. A. Majewski, P. Vogl, The structure of binary compounds edited by F. R. de Boer and D. G. Pettifor, Elsevier, Amsterdam, 1989.
5. T.B. Boykin, J.P.A. van der Wagt, and J.S. Harris, Phys. Rev. **B43**, 4777 (1991).
6. A. Di Carlo, P. Lugli, Semicon. Sci. Technol. **10**, 1673 (1995).
7. J. N. Schulman, Y. C. Chang, Phys. Rev. **B31**, 2056 (1985).
8. A. Zunger, C.-Y. Yeh, L.-W. Wang, S. B. Zang, Proceedings ICPS-22, 1763 (1994).
9. A. Di Carlo, S. Pescetelli, M. Paciotti, P. Lugli, M. Graf, Solid State Communications **98**, 803 (1996)
10. A. Di Carlo, P. Vogl, and W. Pötz, Phys. Rev. **B50**, 8358 (1994).
11. S. Froyen, Phys. Rev. **B39**, 3168 (1989).
12. E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, D. Sorensen, LAPACK User's Guide, SIAM, Philadelphia, 1992.
13. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, Numerical recipes, Cambrige University Press, 1986.
14. M. Graf and P. Vogl, Phys. Rev. **B51**, 4940 (1995).
15. C. Priester, G. Allan, M. Lannoo, Phys. Rev B **37**, 8519 (1988); **38**, 9870 (1988); **38**, 13451 (1988); Y. Foulon, C. Priester, Phys. Rev. B **44** 5889 (1991).
16. F. Aniel, P. Boucaud, A. Sylvestre, P. Crozat, F. H. Julien, R. Adde, and Y. Jin, J. App. Phys. **77**, 2184 (1995): N. Shigekawa, T. Enoki, T. Furuta, and H. Ito
17. A. Reale, A. Di Carlo, and P. Lugli, VLSI-design, in press (1997)
18. A. Reale, A. Di Carlo, and P. Lugli, In Proc. V-European Gallium Arsenide and related III-V compound: Application symposium (GAAS 97) Bologna, Italy, 1997, pag. 305

TIGHT-BINDING FORMALISM FOR IONIC FULLERIDES AND ITS APPLICATION TO ALKALI-C$_{60}$ POLYMERS

SUSUMU SAITO*, STEVEN G. LOUIE**, MARVIN L. COHEN**

*Department of Physics, Tokyo Institute of Technology
2-12-1 Oh-okayama, Meguro-ku, Tokyo 152, JAPAN
**Department of Physics, University of California at Berkeley, and
Materials Sciences Division, Lawrence Berkeley National Laboratory
Berkeley, California 94720

ABSTRACT

We present a tight-binding formalism which can properly treat various ionic fullerides. In the Hamiltonian we include the intrafullerene Coulomb repulsion energy and the Madelung energy of the ionic lattice, both of which depend on the possible charge disproportion between fullerenes. This Hamiltonian requires a self-consistent treatment, but it is applicable to much larger systems than first-principles methods. Using this formalism we have studied the electronic structure of the one-dimensional A$_1$C$_{60}$ polymer. The present generalization of the tight-binding model is found to be important for ionic fullerides and a moderate-amplitude charge-density-wave state is found to be a possible stable state.

INTRODUCTION

Materials consist of ionic cores and valence electrons which mutually interact to attain the lowest free-energy state under a given external condition. To predict this microscopic optimum state, which nature automatically finds, usually requires elaborate first-principles quantum-mechanical calculations. This can be done for infinite crystalline systems as well as for finite systems having a reasonable number of independent degrees of geometrical freedom.

On the other hand, the recent macroscopic production of cage-network all-carbon clusters called "fullerenes" such as C$_{60}$, C$_{70}$, and C$_{84}$, and also the discovery of the carbon nanotubes opened a new field in materials science and engineering. These systems have a variety of geometries with a large number of constituent atoms in a unit cell or in a molecular unit. Due to their large unit size and to experimental difficulties in determining geometries of carbon-based complex materials because of the weak X-ray scattering ability of light elements, theoretical structure studies represent an important tool in this field. However, first-principles calculations can be computationally highly expensive, therefore, less expensive theoretical methods such as interatomic model potentials and tight-binding methods are now attracting renewed interest [1]. In particular the tight-binding model constructed so as to reproduce first-principles results for various carbon systems has been proven to be useful in studying geometries and the electronic structure of these all-carbon systems [2,3].

Among various interesting subfields of this growing field of new carbon materials, the research on the alkali-doped fullerides (A$_N$C$_{60}$, A=Na, K, Rb, and/or Cs) have been at-

tracting great interest since it has resulted in various interesting material phases including superconductivity [4]. Also recently, several A_1C_{60} fullerides are found to have stable polymerized phases [5], which now form another new class of carbon-based crystalline materials consisting of both sp^2 and sp^3-hybridized C atoms. Although the charge transfer from alkalimetal atoms to the fullerene units usually takes place in these fullerides, the application of the standard all-carbon tight-binding model with an additional number of electrons may not give a correct microscopic optimum state since there remain two important terms to be included for these ionic fullerides. One is the intra-C_{60} Coulomb repulsion energy, and the other is the electrostatic energy of the ionic crystalline lattice (Madelung energy). In the present work, we generalize the tight-binding model by including these two important terms for ionic fullerides. This approach is applied to the K_1C_{60} polymer and these terms are found to be very important. The obtained stable state is a charge-density-wave (CDW) state with a considerably smaller amplitude than the CDW state obtained by the tight-binding model without these two additional terms.

FORMALISM

In the usual tight-binding (TB) model [2], the total energy including the short-range repulsive term E_{rep} is given as

$$E_{tot} = 2\sum_i^N \epsilon_i + E_{rep}(\mathbf{R_1, R_2}, ...),\qquad(1)$$

where N is the number of occupied states and the factor 2 is from the spin degeneracy. The ϵ_is are the eigenvalues of the transfer-matrix Hamiltonian with a basis of atomic $2s$ and $2p$ orbitals of the constituent C atoms. This first term is derived by minimizing the electronic energy

$$E_{ele}^0 = 2\sum_i^N \langle\psi_i|H_{TB}|\psi_i\rangle,\qquad(2)$$

with respect to the expansion coefficients $\{C_i^j\}$ of the electronic state ψ_i by atomic orbitals ϕ_j:

$$\psi_i(\mathbf{r}) = \sum_j C_i^j \phi_j(\mathbf{r}),\qquad(3)$$

under the normalization condition $\langle\psi_i|\psi_i\rangle = 1$. By introducing the Lagrange multipliers ϵ_i, we now minimize

$$I^0 = 2\sum_i^N \langle\psi_i|H_{TB}|\psi_i\rangle - 2\sum_i^N \epsilon_i\langle\psi_i|\psi_i\rangle,\qquad(4)$$

which results in a standard matrix diagonalization calculation.

In the case of ionic fullerides, on the other hand, the electrostatic long-range force, i.e. the Madelung energy (E_M), should be considered in addition to the short-range repulsive interatomic force. Also the intrafullerene Coulomb repulsion energy (U) should be considered. This generalization of the TB model for ionic fullerides can be done in the following

way. We consider the electronic energy of the from:

$$E_{ele} = 2\sum_{i}^{N} \langle \psi_i | H_{\mathrm{TB}} | \psi_i \rangle + \sum_{k} \frac{1}{2} U (n_k - 240)^2 + E_M(n_1, n_2, ...),\qquad(5)$$

where 240 is the number of valence electrons in a neutral C_{60} unit, the index k runs on C_{60} units, and n_k is the number of total valence electrons on the k-th C_{60} unit;

$$n_k = 2\sum_{i}^{N} \sum_{j \text{ on } k-\text{th } C_{60}} |C_i^j|^2.\qquad(6)$$

Then, the value to be minimized instead of I^0 in (4) as a function of $\{C_i^j\}$ is

$$I = E_{ele} - 2\sum_{i}^{N} \epsilon_i \langle \psi_i | \psi_i \rangle.\qquad(7)$$

The Madelung energy as a function of $\{n_k\}$ is first numerically evaluated by using the Ewald method, and then fitted by a quadratic form, which has been confirmed to be highly accurate for $A_1 C_{60}$ fullerides as will be shown later.

The introduction of the above site-depecent occupation number of C_{60} units (n_k) now enables us to deal with a various kinds of CDW states having electronically, and consequently, geometrically inequivalent C_{60} units. In the case of alkali-doped C_{60} superconductors, the strong electron-phonon interaction generally expected for carbon-based materials is considered to be one of main reasons for their high transition temperatures. Since the CDW state should also be one of the favorable states in the strong electron-phonon interaction systems, the present course of generalization of the TB model should be of importance in various electron-doped fullerides.

Although the present formalism is for ionic crystalline fullerides, the intrafullerene Coulomb repulsion should be also important for studying the stability of ionized isolated fullerenes, C_{60}^{n-}. Without the second term in (5), the first term, the one-body term, would stabilize unrealistic highly-ionized fullerenes.

APPLICATION TO $A_1 C_{60}$ POLYMER

The rock-salt geometry $A_1 C_{60}$ (A=K, Rb, or Cs) fullerides, which are stable at high temperatures, have been found to transform to a one-dimensionally polymerized C_{60} in an orthorhombic phase with two C_{60} units (two chains) per cell [5]. Their transport properties show complicated temperature and alkali-element dependences [6]. It has been discussed experimentally that some kind of phase transition such as a spin-density wave (SDW) or a CDW seems to take place at low temperatures. However, the direct observation of neither SDW nor CDW has been reported. In the present work, we apply the above generalized TB model to study the stability of a possible CDW state and its amplitude.

It is expected that one of the two added terms in our TB model, the intrafullerene Coulomb repulsion, should tend to depress the CDW amplitude. The other term, the Madelung energy, on the other hand, tends to enhance the amplitude. Therefore, a quantitatively reliable evaluation of these two terms is required. In our previous work [7], we

have studied the electronic energy levels of isolated C_{60} and C_{60}^- clusters and nearly constant shift of energy levels upon the ionization has been found. Therefore, the bare intrafullerene Coulomb repulsion, U, should be equal to this shift, 3.4 eV.

The Madelung energy should be, on the other hand, evaluated numerically for each CDW state to be formed. In the quasi one-dimensional chain-like materials, the density-wave formation is usually observed not only along the chain axis but also along the remaining two directions. Therefore, in the present work, the CDW state with the unit cell of $a' = 2 \times a$, $b' = 2 \times b$, and $c' = 2 \times c$ is considered. Here, the a axis is the polymerization axis and the experimentally observed lattice constants for K_1C_{60} (a=9.109, b=9.953, and c=14.321 (Å)) are used. Hence, the value E_M can be evaluated as a function of the CDW amplitude, i.e., the population difference between two kinds of C_{60} units, δ; n_k is either 241+δ or 241-δ. Here, the more and less populated C_{60} units are labeled by A and B, respectively. The electronic energy per one A-B pair is then given as

$$E_{ele} = 2\sum_i^N \langle \psi_i | H_{\mathrm{TB}} | \psi_i \rangle + U(1 + \delta^2) + E_M(\delta). \tag{8}$$

For the present CDW state, the following expression,

$$E_M(\delta) = c_0 + c_2\delta^2, \tag{9}$$

with c_0=$E_M(0)$=-7.313 and c_2=-2.474 (eV) is found to reproduce the numerically obtained value for K_1C_{60} within an error of less than 0.001% for the whole range of δ (0~1). Then, it is straightforward to show that, to minimize I of (7), the coefficients $\{C_i^j\}$ should satisfy

$$\sum_j t_{lj}C_i^j + \tilde{U}(1 \pm \delta)C_i^l = \epsilon_i C_i^l. \tag{10}$$

Here, t_{lj} is the transfer integral between ϕ_l and ϕ_j, $\tilde{U} = U + c_2$, and the upper (lower) case is for the orbital ϕ_l on the $A(B)$-type C_{60}. Since δ depends on $\{C_i^j\}$ to be determined, (10) should be solved self-consistently. Still this iterative calculation is much less expensive than first-principles calculations, and is easily applied to the present CDW state considered.

The self-consistent calculation for K_1C_{60} with a simultaneous geometry optimization by minimizing E_{tot}=E_{ele}+E_{rep} is found to give δ=0.4. On the other hand, the TB model without two additional energies considered in the present work is found to give δ=0.9. This difference clearly indicates the importance of the present generalization of the TB model for ionic fullerides.

In addition to K_1C_{60}, the Madelung energy calculation for Rb_1C_{60} with the same kind of the CDW state has been done and c_2 is found to be -2.456 (eV), which is similar to the value for K_1C_{60}. Hence, the similar CDW amplitude is expected for this material.

CONCLUDING REMARKS

In the case of one-dimensionally polymerized A_1C_{60}, the CDW state, found to be one of the possible low-energy states in the present work, would be further stabilized if the off-center displacement of the alkali ions in a large interstitial site of C_{60} lattice takes place. A

study of this point is now in progress. Also a spin-polarized study of A_1C_{60} with the present TB model would be a very interesting future project for clarifying the relative stability of the CDW and the SDW states in this material.

ACKNOWLEDGMENTS

REFERENCES

1. S. Saito, S. Sawada, and N. Hamada, Phys. Rev. B **45**, 13845 (1992).

2. C. H. Xu, C. Z. Wang, C. T. Chan, and K. M. Ho, J. Phys. Condensed Matter 4, 6047 (1992).

3. B. L. Zhang, C. Z. Wang, and K. M. Ho, J. Chem. Phys. **96**, 7183 (1992).

4. A. F. Hebard, M. J. Rosseinsky, R. C. Haddon, D. W. Murphy, S. H. Glarum, T. T. M. Palstra, A. P. Ramirez, and A. R. Kortan, Nature **350**, 600 (1991); K. Tanigaki, T. W. Ebbesen, S. Saito, J. Mizuki, J. S. Tsai, Y. Kubo, and S. Kuroshima, Nature **352**, 222 (1991).

5. P. W. Stephens, G. Bortel, G. Faigel, M. Tegze, A. Janossy, S. Pekker, G. Oszlanyl, and L. Forro, Nature **370**, 636 (1994).

6. See, for example, F. Bommeli, L. Degiorgi, P. Wachter, O. Legeza, A. Janossy, G. Oszlanyi, O. Chauvet, and L. Forro, Phys. Rev. B **51**, 14794 (1995); J. Hone, M. S. Fuhrer, K. Khazeni, and A. Zettl, Phys. Rev. B **52**, R8700 (1995).

7. S. Saito, in *Clusters and Cluster-Assembled Materials* (edited by R. S. Averback, D. L. Nelson, and J. Bernholc), Proc. 1990 Mat. Res. Soc. Fall Meeting, Proc. Vol.206 (MRS 1991), P.115

ELECTRONIC STRUCTURE, PRESSURE DEPENDENCE AND OPTICAL PROPERTIES OF FeS$_2$

D. NGUYEN-MANH *, D.G. PETTIFOR *, H.M. SITHOLE **, P.E. NGOEPE **, C. ARCANGELI ***, R. TANK ***, O. JEPSEN ***
*Department of Materials, University of Oxford, Parks Road, Oxford OX1 3PH, U.K.
** Materials Modelling Centre, University of the North, Private Bag X 1106, Sovenga, 0727, South Africa.
*** Max-Plank-Institut für Festkörperforschung, Heissenbergstrasse 1, D-70569, Stuttgart, Germany.

ABSTRACT

A revisited electronic structure study of iron pyrite, FeS$_2$, has been performed using a new Tight-Binding Linear Muffin-Tin Orbital (TB-LMTO) technique in which the radii of overlapping MT spheres are determined from a full potential construction. The interstitial spheres were chosen to provide an efficient packing of space while ensuring that the overlap between the spheres remain small. We have found that this treatment of interstitial spheres results in a dramatic improvement in the description of the electronic structure and the binding energy curves for FeS$_2$ in comparison with a previous LMTO calculation. In particular, the energy band gap, the equilibrium lattice constant and the bulk modulus are all in much better agreement with experimental observations. Moreover, the calculated equation of state is in excellent accord with recent measured P-V data up to pressures of 15GPa with overall deviations of less than 10%. The predicted reflectivity spectrum of FeS$_2$ as a function of pressure gives the observed behaviour of the optical edge. The bonding behaviour the orthorhombic marcasite phase of FeS$_2$ is also discussed within this new TB-LMTO formalism.

INTRODUCTION

Transition metal sulfides are a major group of minerals. The iron disulphides and other pyrites are the most widely occuring of the sulphides being found not only in ore deposits but also as accessory minerals in many common rocks [1]. The mining industry is interested in their electronic, magnetic, optical, structural and thermodynamic properties since there are ore formation, mineral processing and environmental mineralogy. A previous electronic structure study of the pyrite-type disulphides (MS$_2$, where M=Mn, Fe, Co, Ni, Cu, Zn) within the local density approximation (LDA) using the Liniear Muffin-Tin Orbital method in the Atomic Sphere Approximation (LMTO-ASA) showed a qualitative relationship between the calculated local density of states and the experimental X-ray photoelectron spectra (XPS) and the Bremsstrahlung Isochromat spectra (BIS) [2]. However, the predicted equation of state as well as the corresponding value of the bulk modulus were poorly reproduced in this paper. In particular, for FeS$_2$, the predicted bulk modulus of 675 GPa was is too high by a factor of 4.6 compared to experiment. These results may give the impression that the LMTO method is not good enough to use for extracting quantitative information about pyrite-type materials. The purpose of this paper is to make a revisited calculation for FeS$_2$ by using a new TB-LMTO technique in which the muffin-tin radii of overlapping spheres is determined from a full potential construction [3]. Our results show that a correct treatment of the interstitial spheres in the pyrite structure removes the previous theoretical disagreement with experimental values, so that the TB-

LMTO method provide a simple and efficient way to predict quantitatively the electronic structure, equation of state, optical properties and bonding character in the pyrite-type materials.

TB-LMTO COMPUTATIONAL THEORY

In the present study the electronic structure was calculated self-consistently using the TB-LMTO technique the details of which have been described in [4]. We use the scalar-relativistic version in the atomic-sphere-approximation including the combined correction (ASA+CC). Exchange and correlation contributions to both atomic and crystal potentials were included through the local-density-functional description using the von Barth-Hedin formula [5]. We also performed the total-energy calculation with the gradient exchange correlation potential using a Langreth-Mehl-Hu non-local correction to the LDA [5]. In the ASA+CC, the one electron potential entering the Schrödinger equation is a superposition of overlapping spherical potential wells with position $\mathbf{R}$ and radii $s_{\mathbf{R}}$, plus a kinetic-energy error proportional to the 4th power of the relative sphere overlap [6]:

$$\omega_{\mathbf{RR'}} \equiv \frac{s_{\mathbf{R}} + s_{\mathbf{R'}}}{|\mathbf{R} - \mathbf{R'}|} - 1 \tag{1}$$

In the pyrite structure and with only atom-centered spheres, the ASA would cause substantial errors, either due to large overlap and misrepresentation of the potential, or due to neglect of charge in the van der Waals gap. Therefore, it is necessary to pack the van der Waals gap with interstitial spheres. In general, the requirements for choosing sphere positions and radii are that the superposition of the spherical potentials approximate the full three-dimentional potential as accurately as possible, so that the overlap error for the kinetic energy be acceptable. Here, as a model for the full potential we use the superposition of neutral-atom potentials and for simplicity, take only the Hatree part. The atomic-centered spheres are then determined by tracing the potential along the lines connecting nearest-neighbour atoms and finding the saddle-points. For a given atom with position $\mathbf{R}$, the distance to the closest saddle-point is taken as the radius of a sphere and usually touch the sphere constructed in the same way for other atoms. The ASA radii are then obtained by inflating these atom-centered non-overlapping spheres until they either fill space or until their overlap $\omega_{\mathbf{RR'}}$ reaches a maximum of 16%. In the latter case, the potential between the atomic potentials must be represented by *additional interstitial spheres*, which are usually repulsive. The positions of these interstitial spheres are first chosen *among the non- occupied symmetry positions of the space group*. Then their radii are chosen in such way that the maximum overlap between an atomic and an interstitial sphere is 18% and maximum overlap between two interstitial spheres is 20%. This precedure has been made automatic in computer program of ref. [4].

FeS$_2$ crystallizes in the cubic primitive structure with space group Pa$\bar{3}$. The experimental value of lattice parameter is a=5.4179 $\mathring{A}$ and the positions of the atoms are generated by Fe: (4a) (0,0,0) and S: (8c) (u,u,u) with u=0.384 [7]. Our Full-Potential LMTO calculation shows that the theoretical value of the internal parameter u is only 0.5 % smaller than experimental. Following the above procedure for FeS$_2$ additional interstial spheres are placed at E: (24d) (x,y,z) with x=0.31681, y=0.09448 and z=0.21362. The muffin-tin sphere radii are s_{Fe}=1.1714 $\mathring{A}$, s_S=1.0864 $\mathring{A}$ and s_E=0.7363 $\mathring{A}$ and with radii, the sum of the MT sphere volumes equals the unit cell volume. At this point we can see a crucial difference in our treatment of interstitial spheres in comparison with the previous calculation

[2] where only 2 empty spheres per formula unit formula had been introduced. It is not clear from their paper [2] where the exact positions of empty spheres are, but with a unit cell of 4 formula units they used only 8 interstial spheres in comparisions to our 24. We think that our introduction of interstitial spheres is physically transparent and, therefore, results in the quantitatively better description of the equation of state for the pyrite structure in general and for FeS_2 in particular. The basis set consisted of Fe 4s, 4p, 3d; S 3s, 3p, 3d and the interstitial sphere E 1s 2p so that in total there are 204 basis functions per unit cell in comparision with only 140 in [2]. All **k**-space integrations were performed by the tetrahedron method. Convergence to self-consistency was achieved with a grid of 11 irreducible **k** points, although subsequently 119 irreducible **k** points were used to generate the band structure and density of states for FeS_2. In all calculations, the tolerance in total energy before self-consistency was deemed to have been achieved was 10^{-2} mRy.

RESULTS

Equation of state

We have performed both LDA and gradient corrections to LDA for calculating the equation of state. Spin-polarised calculations for both ferromagnetic and antiferromagnetic states show that there is no magnetic moment in pyrite FeS_2 in the vicinity of equilibrium. The LDA total energy reaches a minimum at $V_{LDA}=0.97\ V_{exp}$. The equation of state for FeS_2 is analysed by using the Birch-Murhaghan formula and the pressure dependence as a function of volume renormalised by the equilibrium value, V_{eq}, is shown in Fig. 1.
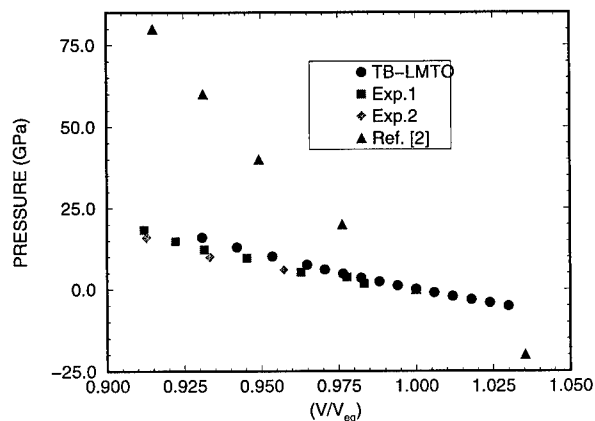
Figure 1. Pressure vs. renormalised volume by equilibrium value

Our TB-LMTO calculations are in very good agreement with two experimental measurements performed recently by Jephcoat: a) hydrostatic (isotropic) compressions in neon medium (Exp.1) and b) nonhydrostatic (uniaxial loading) compressions with no pressure medium (Exp.2) [8]. For instance, the pressure at $V=0.975\ V_{eq}$ is 4.82 GPa which was measured experimentally as 4.28 GPa. Comparing our results with the theoretical calculations in [2], we can see that the new treatment of the empty spheres improves dramatically overall

pressure-volume dependence. Our calculated bulk modulus at the theoretical equlibrium volume is 187 GPa which is within 30% of the observed value of 147 GPa. The non-local correction to the LDA is found to give marginally better results: $V_{nlc}=0.983\ V_{exp}$, $B_{nlc}=182$ GPa and $P_{nlc}(V=0.93\ V_{eq})=15.2$ GPa in comparison with 15.9 GPa within the LDA calculation.

Electronic structure

The calculated total and partial electronic densities of states (DOS) for pyrite FeS_2 is presented in Fig.2. The valence band is fully occupied, the top of the valence band setting the energy zero. Our calculations indicate that FeS_2 is a semiconductor with an *indirect band gap* which is coherent with a recent density-functional calculation using norm-conserving pseudopotentials [9]. However, our predicted LDA and gradient corrected band gap E_g take the values of 0.71 eV and 0.6 eV, respectively, comparing with 0.52 eV (LDA) and 0.3 eV (GGA) calculated in [9]. Experimental values of E_g show considerable dispersion (from ~ 0.7 eV to ~ 2.6 eV) as has been recently reviewed [10]. Optical absorption edge measurements of pyrite indicate a value of 0.95 eV for an indirect gap transition [11].
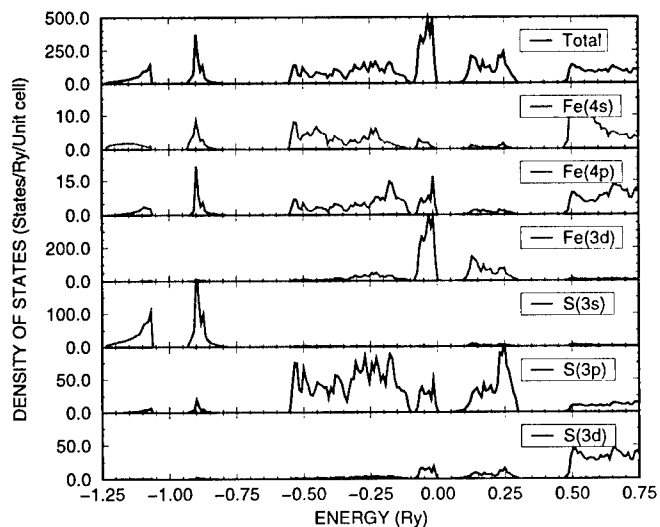
Figure 2. Calculated total and partial DOS for FeS2-pyrite

It is important to emphasise that although there are contributions to the total DOS at the valence and conduction bands coming respectively from the t_{2g} and e_g symmetries of the Fe(3d) orbitals [12], it is clear from Fig. 2 that *physical nature of the band gap is due to strong bonding between Fe(3d) and S(3p) states*. In fact, apart from the isolated s-band at high binding energy ($\sim$ -1 Ry), the valence band between -0.5 Ry and zero energy and the conduction band (up to 0.25 Ry) originate essentially from the Fe(3d)-S(3p) hybridisation states as has already been well documented for the case of $RuAl_2$ and $RuGa_2$ [13].

404

<u>Optical properties</u>

The refelectivity of FeS_2 is calculated from the frequency dependent conductivity tensor $\sigma_{xx}(\Omega)$ via the complex optical constant, κ= n+ik, by the relation

$$R(\Omega) = \frac{(n(\Omega) - 1)^2 + k^2(\Omega)}{(n(\Omega) + 1)^2 + k^2(\Omega)}, \tag{2}$$

where Ω describes the photon energy. The results, which we obtained using the ESOCS programme from MSI, are shown in Fig. 3 for different values of the calculated pressure.
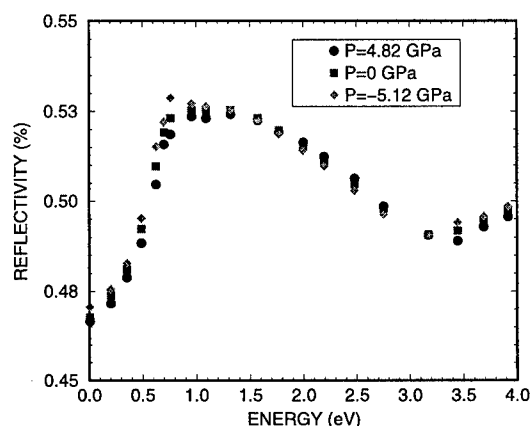
Figure 3. Reflectance spectra of FeS_2 for different pressures

We find that *under high pressure the reflectivity peak is shifted to higher energy levels in FeS_2* which is consistent with measurements at 0.2 GPa and 4.2 GPa in [14]. Our calculated maximum refelectivity is larger than experimental [14] by approximately 23 %, although it does agree with more recent measurements by Ferrer et.al. [10]. The minimum in the reflectivity around 3 eV reflects the conduction band width just above the semiconductor gap as shown in the DOS from Fig. 2.

DISCUSSION

In order to understand the bonding properties in FeS_2, we have performed additional electronic structure calculations for the orthorhombic marcasite phase (space group Pnnm), the second form of dimorphic FeS_2, by using the same TB-LMTO scheme. We predict semi-conductor behaviour with an indirect band gap E_g of 0.85 eV which is twice as large as the TB calculation (0.4 eV) from [12]. Interestingly this value is also larger than our calculated for the pyrite phase of FeS_2. As far as we know, there are no experimental data of optical and conductivity gaps in the marcasite phase of FeS_2. However, whereas the S-S bond length is 1.6 % longer in pyrite than in marcasite, *the Fe-S bond length in marcasite is about 1.7 % shorter* [15]. The larger theoretical gap predicted in the marca-site phase probably originates from the stronger 3d-3p hybridisation reflected in the Fe-S

bonding behaviour discussed above. This is further evidence that the covalent character of the bonding plays a key role in formation of the semiconductor gap in FeS_2 rather than the conventional picture of the crystal-field splitting within transition metal d-band [12]. Assuming fixed internal coordinates, we find the calculated ratio of equilibrium volume between the two phases, $(V_{eq}^{pyr}/V_{eq}^{mar})$, of 0.957 which compares well with the experimental value of 0.974. We also found a small total energy difference between these polymorph phases of 9.5 kJ/mol which overestimates (within the LDA calculations) the experimental values of the enthalpy difference of 4-5 kJ/mol [16].

ACKNOWLEDGMENTS

REFERENCES

1. D.J. Vaughan and A.R. Lennie, Sci. Progress Edinburgh, **75**, 371, (1991).

2. W.M. Temmerman et.al., Phys. Chem. Minerals, **20**, 248, (1993).

3. O.K. Andersen, O. Jepsen and G. Krier, in Method of Electronic Structure Calculations, eds. V. Kumar, O.K. Andersen and A. Mookerjee, Worth Scientific, p.1, (1994).

4. G. Krier, O. Jepsen, A. Burkhart and O.K. Andersen, The TB-LMTO-ASA program, Stuttgart, April (1995).

5. U. von Barth and L. Hedin, J. Phys. C, **5**, 1629, (1972); D.C. Langreth and M.J. Mehl, Phys. Rev. B, **28**, 1809, (1983); C.D. Hu and D.C. Langreth, Phys. Rev. B, **33**, 943, (1986).

6. O.K. Andersen, et.al., Mater. Res. Symp. Proc., **253**, 37, (MRS 1992).

7. In Pearson's Handbook of Crystallographic Data for Intermetallic Phases, eds. P. Villars and L.D. Calvert, Amer. Society for Metal, vol. 3, 2220, (1985).

8. A.P. Jephcoat, to be published.

9. Y. Zeng and N.A.W. Holzwarth, Phys. Rev. B **50**, 8214, (1994).

10. I. J. Ferrer et.al., Solid State Commun., **74**, 913, (1990).

11. M.S. Seehra and S.S Seehra, Phys. Rev. B, **19**, 6620, (1979).

12. D.W. Bullet, J. Phys. C, **15**, 6163, (1982).

13. D. Nguyen-Manh et. al., Solid State Commun., **82**, 329, (1992).

14. N. Mori and H. Takahashi, in High-Pressure Research in Mineral Physics, eds. M.H. Manghnani and Y. Syono, American Geophysical Union, 341, (1987).

15. C. Sourisseau, R. Cavagnat and F. Fouassier, J. Phys. Chem. Solids, **52**, 537, (1991).

16. A.R. Lennie and D.J. Vaughan, Amer. Mineralogist, **77**, 1166, (1992).

Effects of Grain-Boundaries in Superconducting Materials

J.J. Hogan-O'Neill, A.M. Martin and James F. Annett, University of Bristol, H.H. Wills Physics Laboratory, Royal Fort, Tyndall Ave, Bristol BS8 1TL, United Kingdom.

ABSTRACT

We examine the effects of grain-boundaries on the order-parameter and critical-currents in superconductors. We use a geometrical model of the lattice structure of grain-boundaries. We solve the Bogoliubov-de Gennes equation using the Recursion Method to obtain the self-consistent BCS gap function Δ and the local density of states $N(E)$ near the boundary. Imposing a phase difference across the boundary we calculate the supercurrent and hence obtain the critical-current, I_c, of the junction.

INTRODUCTION

Since the discovery of high T_c superconductors over 10 years ago it has been generally agreed there is an unconventional symmetry of the order parameter, the most likely being $d_{x^2-y^2}$ pairing. Indeed, this can be shown from several different experiments [1]. However, due to their highly complex crystal structure, the difficulty of fabricating single crystals without twinning, and the orthorhombic nature of the CuO_2 planes, it has been anticipated by some that there is a mixing of s-wave (local) and d-wave (non-local) symmetries [2]. Tseui *et al* [3] have reported a half-integer flux quantisation through a superconducting ring with an odd number of grain-boundary weak-links in the ring. They attribute this to an unconventional pairing state, possibly $d_{x^2-y^2}$.

One way to help clarify the role of grain-boundaries and twinning is to model interfaces between s-wave and d-wave superconductors. All calculations up until now have modelled perfectly flat interfaces and have assumed a specific form for the order parameter [4, 5, 6]. Martin and Annett [7] have taken the more realistic step of inputing an idealised order-parameter but allowing it to evolve through a self-consistent algorithm to solve the Bogoliubov-de Gennes equations. Annett and Goldenfeld [8], and Litak *et al* [9], have already undertaken similar calculations but only for a purely local-interaction corresponding to s-wave pairing. The Greens functions are calculated using the Recursion Method [10] and thus it is possible to determine the local and non-local charge densities, order-parameters, and also, the local quasi-particle density of states.

This work also uses that algorithm, but instead of using a perfectly flat interface we now use an idealised grain-boundary (figure 1) with periodicity and reflective-symmetry around the line $x=0$. This periodicity allows us to select a few sites for self-consistent calculations which can then be mirrored onto similar sites in the rest of the sample: these inequivalent sites are highlighted in figure 1. It is also possible to calculate the supercurrent across the boundary with differing phases of the order-parameter imposed on either side in the bulk. From this we obtain the critical current for the junction. By doing this for various grain-boundary angles it will be possible to determine how the critical current depends on the grain boundary angle. In this paper we shall discuss the s-wave scenario; the d-wave case shall be reported elsewhere [11].
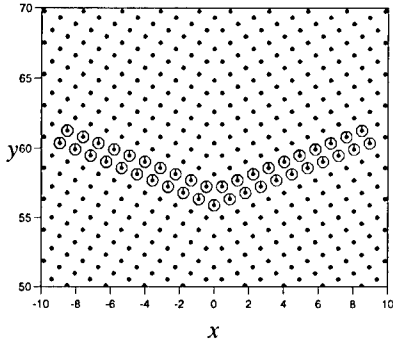
Figure 1: *Construction of the model grain-boundary (symmetric). By periodicity, it is possible to carry out self-consistent calculations on a few sites (highlighted) which are then mirrored onto similar sites in the rest of the sample.*

THEORY

The equation under consideration is the Bogoliubov-de Gennes Matrix Equation on a tight binding lattice:

$$\sum_j \mathbf{H_{ij}} \begin{pmatrix} u_j^n \\ v_j^n \end{pmatrix} = \mathbf{E_n} \begin{pmatrix} u_i^n \\ v_i^n \end{pmatrix} \tag{1}$$

with

$$\mathbf{H_{ij}} = \begin{pmatrix} H & \Delta_{ij} \\ \Delta_{ij}^\star & -H^\star \end{pmatrix} \tag{2}$$

and

$$H = \begin{cases} H_{ii} = (\epsilon_i - \mu + \frac{1}{2}U_{ii}n_{ii}) & (local) \\ H_{ij} = (t_{ij} + \frac{1}{2}U_{ij}n_{ij}) & (non-local) \end{cases} \tag{3}$$

where μ is the chemical potential, ϵ_i is the normal on-site energy of site i and t_{ij} is the hopping integral between site i and site j, u_i^n and v_i^n are the particle and hole amplitudes on site i associated with an eigenenergy E_n. We model t_{ij} to be of the form

$$t_{ij} = \frac{\sqrt{2} - r_{ij}}{\sqrt{2} - 1}, \quad 0 \le r_{ij} \le \sqrt{2} \tag{4}$$

where r_{ij} is the distance, between sites i and j, in units of the lattice constant. The on-site and off-site interaction terms $\frac{1}{2}U_{ii}n_{ii}$ and $\frac{1}{2}U_{ij}n_{ij}$ are the Hartree-Fock potentials corresponding to the on-site interaction U_i and the non-local interaction U_{ij}. The quantities of

interest are the local charge-density, n_{ii}, and order-parameter Δ_{ii}. Writing these quantities in a Greens function formalism we obtain

$$n_{ii} = \sum_{\sigma} \langle c^{\dagger}_{i\sigma} c_{i\sigma} \rangle = \frac{1}{\pi} \int_{-E_c}^{E_c} (G_{ii}^{++}(E + \imath\eta) - G_{ii}^{++}(E - \imath\eta)) f(E) dE, \tag{5}$$

$$\Delta_{ii} = U_{ii} \langle c_{i\uparrow} c_{i\downarrow} \rangle = \frac{1}{2\pi} U_{ii} \int_{-E_c}^{E_c} (G_{ii}^{+-}(E + \imath\eta) - G_{ii}^{+-}(E - \imath\eta))(1 - 2f(E)) dE, \tag{6}$$

$$n_{ij} = \sum_{\sigma} \langle c^{\dagger}_{i\sigma} c_{j\sigma} \rangle = \frac{1}{\pi} \int_{-E_c}^{E_c} (G_{ij}^{++}(E + \imath\eta) - G_{ij}^{++}(E - \imath\eta)) f(E) dE, \tag{7}$$

$$\Delta_{ij} = U_{ij} \langle c_{i\uparrow} c_{j\downarrow} \rangle = \frac{1}{2\pi} U_{ij} \int_{-E_c}^{E_c} (G_{ij}^{+-}(E + \imath\eta) - G_{ij}^{+-}(E - \imath\eta))(1 - f(E)) dE. \tag{8}$$

where the $+/-$ subscripts on the Greens functions refer to particles and holes respectively, and E_c is the cut-off energy ($\hbar\omega_D$ in conventional superconductors). These expressions arise by making the Bogoliubov Transformation for the fermion operators and then by realising that the terms in $u_i v_j^*$ are the residues of the Greens function for the previously stated Bogoliubov-de Gennes Hamiltonian [12]. By solving the above equations in conjunction with equation (1) we can iterate to a self-consistent solution.

In order to calculate supercurrents across the boundary we impose a phase-difference across the junction. The supercurrent, I_{ij}, between the sites i and j is derived by making the Peierls substitution for the overlap integral t_{ij}

$$t_{ij} \mapsto t_{ij} \exp\left[-i \left(\frac{e}{\hbar} A_{ij} + \phi_i - \phi_j\right)\right] \tag{9}$$

where A_{ij} is the line-integral of the vector potential along the link i-j and $\phi_i = \arg(\Delta_{ii})$. Then using

$$I_{ij} = \frac{\partial H}{\partial A_{ij}} \tag{10}$$

yields

$$I_{ij} = \frac{t_{ij} e}{\hbar} \Im \left[\frac{1}{i\pi} \int_{-\infty}^{+\infty} \left[G_{ij}^{++}(E + \imath\eta) - G_{ij}^{++}(E - \imath\eta) \right] e^{-i\theta_{ij}} f(E) dE \right] \tag{11}$$

where $\theta_{ij} = \frac{e}{\hbar} A_{ij} + \phi_i - \phi_j$, and $f(E)$ is the Fermi-Dirac distribution. Note that even for an s-wave order-parameter when calculating the current, we use a non-local Greens function. It is also a trivial matter to calculate the local quasi-particle density of states from the well known relation

$$N(E) = -\frac{1}{2\pi} \Im \left[G_{ii}^{++}(E + \imath\eta) - G_{ii}^{++}(E - \imath\eta) \right]. \tag{12}$$

The Recursion Method allows us to calculate all of the necessary Greens functions efficiently, as a continued fraction of 2×2 matrices.

RESULTS

This paper reports calculations undertaken for an s-wave superconductor with a simple model of a geometric grain-boundary. It is hoped that the construction of this model will give a qualitative understanding of both large and small angle scenarios, and where twinning and un-twinned crystals can be analysed (we study the latter here). A typical

409

model boundary is shown in figure 1. Note that it is periodic in the y-direction and can be divided into *cells* as shown. Each *cell* has a percolation site (the central site in figure 1) through which a supercurrent can always flow. As the angle of the grain-boundary is increased there will be zero hopping between sites on either side if their separation is greater than or equal to $a\sqrt{2}$, a being the lattice constant. For smaller grain-boundary angles we remove sites from one side if the distance to an atom on the opposite side is less than $\frac{2}{3}a$.

We report here on results obtained from the grain-boundary in figure 1, the angle being roughly 53°. For the angle chosen there is a non-zero hopping across the boundary for both strips shown in figure 1, not just the central site. When calculating the Greens function we find that when we generate terms in the continued-fraction after a certain number of levels we can extrapolate for many more levels. In the work presented in this paper we have evaluated 75 levels exactly and extrapolated for a further 3000. The sites on which self-consistent calculations have been made are those highlighted on the two lines in figure 1, i.e. 10 sites away from the edge along both lines giving a total of 41 sites on which self-consistent calculations are made. After this number we set all other sites equal to the bulk value calculated for an infinite square lattice. This can be done provided the quantities being evaluated are well behaved before they reach this distance: we have found that 10 sites from the boundary is sufficient. Indeed this distance is much greater than the superconducting coherence length ξ.
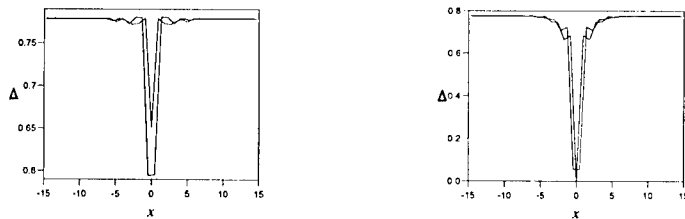
Figures 2(a) and 2(b): *The self-consistent magnitude of the order-parameter as we move across the grain-boundary for a phase-difference of 0° (left) and 180° (right).*

When performing a self-consistent calculation with a local interaction, the only free parameters which need to be defined are U_{ii} and E_c: given these all other parameters are determined self-consistently. For the rest of the paper we will be concerned with the situation where $U_{ii} = -3.5$ and $E_c = 3.0$ in units of the hopping-integral in the bulk.

In figure 2(a) we plot the magnitude of the order-parameter against the x-ordinate, distance from the boundary. We have plotted two lines corresponding to the two lines of inequivalent sites in figure 1. Thus, far away from the boundary we see the value Δ_{ii} tend to a bulk value of around 0.77 in units of the bulk hopping (t_{ij}), but as we approach the boundary we see the characteristic Friedel oscillations [13]. On going through the boundary the order-parameter is depressed by a considerable degree. For smaller angle grain-boundaries (not shown) we see the opposite occur: an enhancement of the order-parameter. This seems to be a general conclusion for this model: a missing site enhances the magnitude of the order-parameter. Note also that this plot is symmetric about $x = 0$ as is expected.

Figure 3 shows the local quasi-particle density of states versus energy for the percolation site (solid line), and the calculated local quasi-paricle density of states in the bulk(dashed line). We observe the characteristic B.C.S. energy-gap around the Fermi Energy: any degree of rounding of the function is due to the existence of a non-zero value for η in the Greens functions. Comparing the dashed and solid lines in figure 4 we can see that the local quasi-particle density of states on the percolation site exhibits much more structure than the bulk local quasi-particle density of states. This is due to the inhomogeneity of the system around the percolation site, which may in turn induce resonances and quasi-localised states at the band-edge.
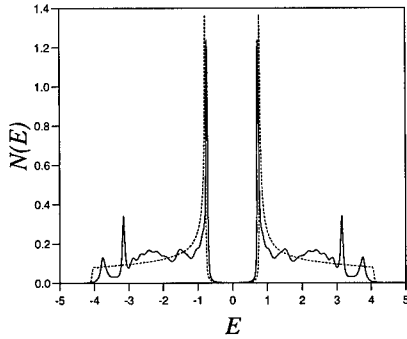
Figure 3: *The two lines depict the local quasi-particle density of states for two different sites in the structure. The solid line is for the percolation site, and the dashed line is for a site deep in the bulk..*

In order to examine supercurrents across the grain-boundary we impose a bulk phase difference $\delta\phi$ across the system. This is done by fixing the order parameter phase, ϕ, deep in the bulk on either side of the junction and allowing it to evolve self-consistently in the region of the grain boundary. In figure 4 we show how the phase (ϕ) varies across a grain-boundary. In this particular case we have set the phase in the bulk to be 0° on one side in the bulk and 30° on the other. Doing this, we can clearly see how the phase naturally evolves on going through the grain-boundary giving a smooth function and therefore a finite phase-gradient. Most previous calculations do not allow the phase-gradient to be determined in a self-consistent manner. However self-consistency is crucial in conserving charge, ie. ensuring that the net current flow into any given site is zero.

In figure 2(b) we have plotted the magnitude of the gap-function $|\Delta(x)|$ in the case where the phase across the junction is 180°. Comparing this with figure 2(a) we can now see that the magnitude of the order-parameter, is not only depressed for this phase difference, but actually goes all the way to zero on the central site. Thus we can see that as one changes the phase-difference across the system, the properties of the grain-boundary are affected, i.e. one cannot just impose a phase difference without considering what effect this may have on the magnitude of the superconducting order-parameter. The magnitude
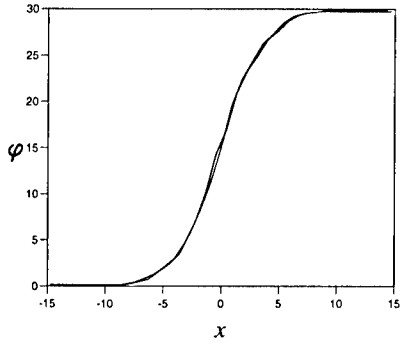
Figure 4: *The self-consistent phase of the order-parameter as we move across the grain-boundary for a bulk phase-difference of* $30°$.

of the order-parameter going to zero is expected in this case since there should be exact cancellation from the two contributions Δ_{ii} from either side of the junction. That this has been reproduced demonstrates the validity of the numerical method employed in this paper.

In figure 5 we have plotted the calculated values for the current across the grain-boundary as a function of the bulk phase difference. The form of this curve is dependent upon the structure of the grain-boundary itself, the sawtooth form for the current phase relation in figure 5 is similar to previous work [14]. We, however, note that the current phase relation must and does obey

$$I(-\phi) = -I(\phi). \tag{13}$$

The important information to gain from figure 5 is not the exact form of this curve but the critical-current (I_c) corresponding to the maximum in figure 5. The critical-current is the maximum value of the current which can pass through the grain-boundary without energy loss. What we have shown in figure 5 is that for a given grain-boundary angle it is possible to calculate this critical-current.

CONCLUSIONS

By solving the Bogoliubov-de Gennes equation using the Recursion Method to calculate the Greens functions we have calculated the local quasi-particle density of states, the magnitude of the order parameter and its phase self-consistently. We have shown how Δ can vary on going through the grain-boundary and also how a change in phase of Δ across the junction affects its value on a given site (e.g. the central site). By imposing

412
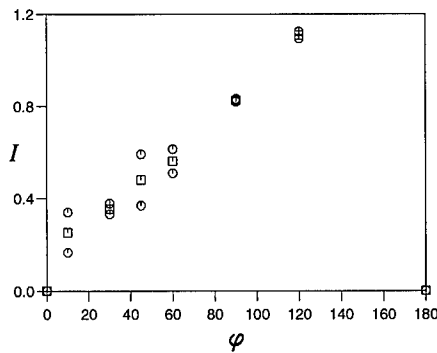
Figure 5: *The total current across the grain-boundary versus the phase-difference. The boxes denote the calculated value, and the circles denote the calculated error.*

a phase difference across the boundary we have also constructed a profile of how this phase-difference affects the value of the supercurrent across the boundary.

By repeating this process for many different grain-boundary orientations it will be possible to see how the critical-current of the boundary varies with the grain-boundary angle, thus enabling us to calculate the critical-current for macroscopic wires with many random grain-boundaries in them. These ideas are now being extended to consider the case for the high temperature superconductors where the order-parameter is likely to have $d_{x^2-y^2}$ symmetry. Other cases of interest shall be to study twinned and orthorhombic scenarios both separately and jointly.

REFERENCES

1. J.F. Annett, N. Goldenfeld, A.J. Leggett in Physical Properties of High Temperature Superconductors, Vol. 5, D.M. Ginsberg (ed.) (World Scientific, Singapore, 1996).

2. A.G. Sun, D.A. Gajewski, M.B. Maple and R.C Dynes, Phys. Rev. Lett. **72**, 2267 (1994).

3. C.C. Tsuei *et al.*, Phys. Rev. Lett. **73**, 593 (1994).

4. Y. Tanaka and S. Kashiwaya, Phys. Rev. Lett. **74**, 3451 (1995).

5. Y. Tanaka, Phys. Rev. Lett. **72**, 3871 (1994).

6. Y. Tanaka and S. Kashiwaya, Phys. Rev. B **56**, 892 (1997).

7. A.M. Martin and J.F. Annett, Cond.-Matt. /9708086.

8. J.F. Annett and N. Goldenfeld, J. Low. Temp. Phys. **89** 197 (1992).

9. G. Litak, P. Miller and B.L. Györffy, Physica C **251**, 263 (1995).

10. R. Haydock, in: Solid state Physics, vol. **35**, eds. Ehrenreich, F. Seitz and D. Turnbull (Academic Press, New York, 1980).

11. J.J. Hogan-O'Neill, A.M. Martin, James F. Annett, To be published.

12. P. Miller PhD Thesis 1994, University of Bristol, U.K.

13. O. Entin-Wohlman and J. Bar-Sagi, Phys Rev. B **18** , 3174 (1977).

14. K.K Likharev, Reviews of Modern Physics **51**, 101 (1979).

Part III

Tight–Binding Simulations

A COMPARISON OF LINEAR SCALING TIGHT BINDING METHODS

A. P. HORSFIELD * D. R. BOWLER**, C. M. GORINGE [†], D. G. PETTIFOR [‡] M. AOKI[§]
*Fujitsu European Centre for Information Technology, 2 Longwalk Road, Stockley Park, Uxbridge, UK
**Physics Department, Keele University, Keele, Staffordshire, ST5 5BG, UK
[†]Australian Key Centre for Microscopy and Microanalysis, Madsen Building (F09), University of Sydney, NSW 2006, Australia. e-mail: chris.goring@cheerful.com
[‡]Oxford University, Department of Materials, Parks Road, Oxford, OX1 3PH
[§]Department of Electrical and Electronic Engineering, Gifu University, 1-1 Yanagido, Gifu, 501-11, Japan

ABSTRACT

Four linear scaling tight binding methods (the density matrix method, bond order potentials, the global density of states method, and the Fermi operator expansion) are described and compared to show relative computational efficiency for a given accuracy. The density matrix method proves to be most efficient for systems with narrow features in their energy gaps, while recursion based moments methods prove to be most efficient for metallic systems.

INTRODUCTION

During computer simulations, the need for accuracy under diverse conditions implies that a quantum mechanical description of interatomic interactions is required. However, the need to handle many atoms implies that computationally efficient implementations are necessary. These two considerations are in competition with one another. A good compromise solution is to use tight binding (TB) which is a simple, but often accurate, quantum mechanical model. While this model can be implemented efficiently on a computer, there is still a bottleneck, namely the cubic scaling with respect to number of atoms of the computational effort required for matrix diagonalisation. This effectively limits the system size to around 100 atoms for molecular dynamics simulations.

Recently, however, a number of schemes have been proposed for which the computational effort scales linearly with number of atoms (so called O(N) methods) [1-11]. In this paper, the density matrix method [1, 2] (DMM), the bond order potential method [6, 7] (BOP), the global density of states method [10] (GDOS), and the Fermi operator expansion method [8] (FOE) are compared for two example systems. The aim of the investigation is to discover which method is computationally most efficient for a given level of accuracy for a given system.

METHODS

The linear scaling methods currently available for TB can be broken down into two broad areas: variational methods and moments based methods. In this section a brief overview of the methods is given according to the two categories.

Density Matrix Method

This is a variational method. This method was proposed simultaneously by Li et $al.$ [1], and by Daw [2], though from different arguments. The method revolves around the density matrix. The number of electrons in the system, the band energy for the ground state and the corresponding contribution to the forces from the band energy can all be written in terms of the density matrix $\hat{\rho}$:

$$N_e = 2\mathrm{Tr}[\hat{\rho}] \tag{1}$$

$$U = 2\mathrm{Tr}[\hat{\rho}\hat{H}] \tag{2}$$

$$\vec{F}_i = -2\mathrm{Tr}\left[\hat{\rho}\frac{\partial \hat{H}}{\partial \vec{R}_i}\right], \tag{3}$$

where Tr indicates taking the trace of a matrix, $\hat{H}$ is the Hamiltonian, and μ is the chemical potential. The ground state energy can be found by minimising U with respect to $\hat{\rho}$ subject to two constraints:

417

idempotency of the density matrix ($\hat{\rho}^2 = \hat{\rho}$, which is equivalent to $\hat{\rho}$ having eigenvalues of 0 and 1) and constant number of particles ($N_e = $ constant). In the density matrix method, the elements of the density matrix are used as variational degrees of freedom with respect to which the energy is minimised. To impose idempotency on the density matrix, it is replaced with the result of the McWeeney transformation [12] of a trial density matrix,

$$\tilde{\rho} = 3\hat{\rho}^2 - 2\hat{\rho}^3, \tag{4}$$

where $\hat{\rho}$ is identified as the *trial* density matrix and $\tilde{\rho}$ is identified as the *physical* density matrix, such that the expectation value of an operator $\hat{A}$ is given by $\text{Tr}[\tilde{\rho}\hat{A}]$.

Without the McWeeney transformation, minimisation of U with respect to the elements of $\hat{\rho}$ would result in $\hat{\rho}$ having eigenvalues of $\pm\infty$; once it has been imposed, the band energy has a single local minimum where $\tilde{\rho}$ is equal to the true density matrix, $\hat{\rho}$, and has eigenvalues of 0 or 1. The runaway solutions can be avoided easily by explicit construction [13].

The second constraint is easily achieved by varying the chemical potential, μ, at each step. In the implementation used in this work, the chemical potential is introduced into the minimisation by working with the grand potential (Ω). Minimising the energy and fixing the number of electrons are then performed in a concerted fashion [13]. The grand potential (Ω) is defined by:

$$\Omega = U - \mu N_e = 2\text{Tr}[\tilde{\rho}(\hat{H} - \mu)], \tag{5}$$

and is minimised with respect to the elements of the trial density matrix, $\hat{\rho}$.

In order to achieve linear scaling, the density matrix method takes advantage of the fact that the elements of the density matrix between two atoms tend to zero as the distance between them tends to infinity. A cutoff radius (R_c) is postulated beyond which all elements of the trial density matrix are set to zero. This leads to a sparse density matrix, which gives linear scaling. However, imposing this cutoff leads to a density matrix which is no longer exactly idempotent, though the McWeeney transformation reduces the error. If the trial density matrix has idempotency errors to first order, then the physical density matrix will have idempotency errors to second order. The transformation forces the eigenvalues of the density matrix which emerges to be clustered about 0 and 1, rather than exactly equal to 0 and 1, as would be the case for an idempotent operator. It should be noted that in the implementation of this method used for this work, the cutoff is defined not by a radius, but by the number of hops away from an atom.

From the point of view of carrying out computer simulations with the DMM, the following features of the method should be noted. The method is variational, leading to excellent forces, and cohesive energies that are bounded from below. Efficiency in molecular dynamics simulations is achieved by carrying forward the density matrix from one time step to the next, thereby greatly accelerating the energy minimization. Where this carry forward is not possible (such as for liquids, for which the change in density matrix between MD time steps is large), the method ceases to be very efficient. This method can be very memory intensive, as the whole density matrix, and the product of the Hamiltonian and the density matrix, must be stored. This limits the applicability of the method, making the study of close packed systems essentially impossible with it. Good scaling can be achieved on a parallel computer.

Moments Methods

Although the details of the three moments methods considered here are quite different, the underlying concept of a moment expansion (either of the density of states or of the Fermi function) is common to them all. Thus by way of introduction to all the methods, we begin with a summary of the key concepts involved in evaluating the density of states from its moments.

The band energy of a system, which is evaluated from the total density of states, can be decomposed into a sum of contributions from the individual atoms by writing $n_{total}(E)$ (the total density of states of the system, and hence a *global* quantity) as a sum over the *local* densities of states [14] of the system:

$$n_{total}(E) = \sum_{i\alpha} n_{i\alpha}(E), \tag{6}$$

where i is a site index, and α is an orbital index.

The density of states is a function with a finite width, a shape and a centre. These properties (and hence the density of states) can be described by the *moments* of the density of states [15]. The pth moment of the local density of states, $n_{i\alpha}(E)$ is $\mu_{i\alpha}^{(p)}$, where

$$\mu_{i\alpha}^{(p)} = \int E^p n_{i\alpha}(E)dE. \tag{7}$$

418

There is a useful identity [16], which states that the pth moment of the density of states projected onto the orbital $|i\alpha\rangle$ equals the pth moment of the Hamiltonian projected onto the same orbital:

$$\mu_{i\alpha}^{(p)} = \int E^p n_{i\alpha}(E)dE = \langle i\alpha|\hat{H}^p|i\alpha\rangle. \tag{8}$$

This identity allows an identification to be made between the pth moment of the density of states and powers of the Hamiltonian matrix:

$$\mu_{i\alpha}^{(p)} = \sum_{j_1\beta_1...j_{p-1}\beta_{p-1}} H_{i\alpha,j_1\beta_1}H_{j_1\beta_1,j_2\beta_2}...H_{j_{p-1}\beta_{p-1},i\alpha}. \tag{9}$$

In turn, the process of multiplying the Hamiltonian matrix into itself can be mapped onto a process of hopping around the lattice along closed paths of length p, by equating $H_{i\alpha,j\beta}$ with a hop between orbital $i\alpha$ and orbital $j\beta$. This then leads to a simple interpretation of local bonding in terms of the local connectivity of atoms.

While moments indeed offer a promising way to reconstruct a function, most possible implementations are numerically unstable. An important exception is the recursion method [17]. This is an optimal Green's function method for building densities of states from moments.

To summarise the above, we note first that the density of states can be reconstructed from its moments. Further the moments can be obtained directly as powers of the Hamiltonian matrix. In turn this allows us to interpret the density of states (and hence the local bonding) in terms of the local connectivity of the atoms, provided we identify the Hamiltonian matrix with a hopping process. It is this local description that leads to the O(N) scaling of these methods. Finally, to turn these appealing intuitive concepts into a stable and rapidly convergent numerical method, we introduce the recursion scheme.

Global Density of States Method

The recursion method is an optimal method for generating the density of states, and hence the band energy, from the moments of the Hamiltonian. For molecular dynamics simulations, however, we also need the atomic forces. These have a contribution from the band energy, which is given by the derivative of the band energy with respect to atomic positions:

$$U_{band} = -\frac{1}{\pi}\sum_{i\alpha}\lim_{\eta\to 0}\text{Im}\int dE G_{i\alpha,i\alpha}(E+i\eta)Ef(x). \tag{10}$$

where $f(x)$ is the Fermi function, $x = (E-\mu)/k_BT$, and $\hat{G}(Z) = [Z - \hat{H}]^{-1}$ is the Green's function. The contribution to the force on atom k from the band energy (at constant chemical potential) is then:

$$\begin{aligned}F_{band}^{(k)} &= -\frac{\partial U_{band}}{\partial \vec{r}_k}\\ &= -\frac{1}{\pi}\sum_{i\alpha}\lim_{\eta\to 0}\text{Im}\int dE\frac{\partial G_{i\alpha,i\alpha}(E+i\eta)}{\partial \vec{r}_k}Ef(x).\end{aligned} \tag{11}$$

Thus we see that the force depends on the derivative of the Green's function. The chain rule for partial differentiation gives:

$$\frac{\partial G_{i\alpha,i\alpha}(Z)}{\partial \vec{r}_k} = \sum_p \frac{\partial G_{i\alpha,i\alpha}(Z)}{\partial \mu_{i\alpha}^{(p)}}\frac{\partial \mu_{i\alpha}^{(p)}}{\partial \vec{r}_k} + \sum_p \frac{\partial G_{i\alpha,i\alpha}(Z)}{\partial \mu_{i\alpha}^{(p)}}\frac{\partial \mu_{i\alpha}^{(p)}}{\partial \vec{r}_k}. \tag{12}$$

The derivatives of the Green's function with respect to the moments can be evaluated straighforwardly. The derivatives of the moments with respect to the atomic positions, however, are more problematic. Formally we can write down the derivative of the pth moment:

$$\begin{aligned}\frac{\partial \mu_{i\alpha}^{(p)}}{\partial \vec{r}_k} &= \frac{\partial}{\partial \vec{r}_k}\sum_{j_1\beta_1...j_{p-1}\beta_{p-1}} H_{i\alpha,j_1\beta_1}H_{j_1\beta_1,j_2\beta_2}...H_{j_{p-1}\beta_{p-1},i\alpha}\\ &= \sum_{j_1\beta_1...j_{p-1}\beta_{p-1}}\left\{\frac{\partial H_{i\alpha,j_1\beta_1}}{\partial \vec{r}_k}H_{j_1\beta_1,j_2\beta_2}...H_{j_{p-1}\beta_{p-1},i\alpha} + ...\right\}.\end{aligned} \tag{13}$$

Unfortunately this expression is, in general, very slow to evaluate on a computer (though it can be used for very low order moment expansions [18]). The reason for this is that the nth moment needs to be evaluated n times for each component of force, and there will be $3N$ components to be considered, where N is the number of atoms in the cluster from which the moment is evaluated.

However, there is a way to greatly accelerate the evaluation of the derivatives, and that is to work with the *global moments*, rather than the local moments. The global moments $(\bar{\mu}^{(n)})$ are defined by:

$$
\begin{aligned}
\bar{\mu}^{(p)} &= \sum_{i\alpha,j_1\beta_1\cdots j_{p-1}\beta_{p-1}} H_{i\alpha,j_1\beta_1} H_{j_1\beta_1,j_2\beta_2} \cdots H_{j_{p-1}\beta_{p-1},i\alpha} \\
&= \sum_{i\alpha} \langle i\alpha \mid \hat{H}^p \mid i\alpha \rangle \\
&= \mathrm{Tr}\left\{\hat{H}^p\right\}.
\end{aligned}
\tag{14}
$$

Because we can permute matrices inside a trace, the derivative of the global moment is given by:

$$
\frac{\partial \bar{\mu}^{(p)}}{\partial \vec{r}_k} = p \sum_{i\alpha,j_1\beta_1\cdots j_{p-1}\beta_{p-1}} \frac{\partial H_{i\alpha,j_1\beta_1}}{\partial \vec{r}_k} H_{j_1\beta_1,j_2\beta_2}\cdots H_{j_{p-1}\beta_{p-1},i\alpha}.
\tag{15}
$$

This is very easy to calculate efficiently on a computer [10], though it can only be evaluated in a stable manner for about the first 20 moments. In the global density of states method, the global moments are used to construct recursion coefficients from which the density of states, band energy and atomic forces are evaluated.

In summary, the use of the global density of states leads to a reduced rate of convergence of the energy with number of moments as compared with the local densities of states for inhomogeneous systems. Further, only about 20 moments can be used before the conversion of moments into recursion coefficients becomes unstable, and since all the moments are stored, this method can be quite memory intensive. This has to be weighed against the chief benefit of using the global density of states, namely that the analytical forces are exact derivatives of the energy. This method can be implemented on a parallel machine with almost perfect scaling, as little information needs to be passed between processors.

Bond Order Potentials

There is an alternative approach to evaluating the forces to that described above. Rather than trying to differentiate the energy exactly, we can appeal to the Hellmann-Feynman theorem [19, 20] (see equation 3). The forces are straightforward to evaluate if we can find the density matrix. The bond order potential method is a scheme for evaluating the density matrix. It should be noted that the Hellmann-Feynman forces will only equal the derivatives of the energy once the density matrix is well converged.

To obtain the forces on the atoms, only the elements of the density matrix within the same range as that of the Hamiltonian matrix are required. These can be evaluated from the off-diagonal elements of the Green's function:

$$
\rho_{i\alpha,j\beta} = -\frac{1}{\pi} \lim_{\eta \to 0} \mathrm{Im} \int dE\, G_{i\alpha,j\beta}(E + i\eta) f(x).
\tag{16}
$$

It has been shown that in general,

$$
\rho_{i\alpha,j\beta} = -\sum_{n=0}^{\infty} \chi^\Lambda_{0n,n0}(\delta a_n^\Lambda)_{i\alpha,j\beta} + 2\sum_{n=1}^{\infty} \chi^\Lambda_{0n,(n-1)0}(\delta b_n^\Lambda)_{i\alpha,j\beta},
\tag{17}
$$

where the response functions for a given total number of electrons, N_e, and electron temperature, T, are defined by

$$
\chi^\Lambda_{0m,n0}(N_e,T) = \frac{1}{\pi} \lim_{\eta \to 0} \mathrm{Im} \int G_{0m}(E + i\eta) G_{n0}(E + i\eta) f(x) dE.
\tag{18}
$$

The terms $(\delta a_n^\Lambda)_{i\alpha,j\beta}$ and $(\delta b_n^\Lambda)_{i\alpha,j\beta}$ are linear combinations of interference terms of the form $\langle i\alpha \mid \hat{H}^n \mid j\beta \rangle$. For more information, see Aoki [7] and Horsfield et al [21])

The key features of this method are: it provides a rapidly convergent expansion for the density matrix in terms of moments; it requires only modest amounts of memory to implement; there are errors in the forces that follow from dependence on the Hellmann-Feynman theorem. This method has been implemented on a parallel machine, and almost perfect scaling is found, as little information needs to be passed between processors.

Fermi Operator Expansion

The Fermi operator expansion is an expansion for the density matrix in terms of powers of the Hamiltonian matrix [22, 23]. It is conceptually the simplest of the moments methods. We can expand the Fermi function, $f(x)$, in powers of E:

$$f(x) = \sum_{p=0}^{\infty} C_p E^p. \tag{19}$$

We can then write (using $\hat{\rho} = f((\hat{H} - \mu)/k_B T)$):

$$\rho_{i\alpha,j\beta} = \sum_{p=0}^{\infty} C_p \langle i\alpha \mid \hat{H}^p \mid j\beta \rangle. \tag{20}$$

From this we can see the following: a polynomial expansion for the Fermi function in powers of E corresponds to an expansion for the density matrix in terms of the moments of the Hamiltonian; the polynomial expansion only need apply over the range of the eigenvalues.

The prescription for this method is then as follows. A polynomial is fit to the Fermi function in powers of E. A polynomial of order $1.5W/k_B T$, where W is the bandwidth (the difference between the highest and the lowest eigenvalues), is found to give well converged results. The density matrix is then constructed from the moments of the Hamiltonian matrix.

If the simple moments of the Hamiltonian are used, the method very quickly becomes unstable. Thus in practise Chebyshev moments are used [22, 23].

It should be noted that this method *requires* the use of a finite electron temperature: an infinite number of moments would be needed to reach $T = 0$. Since the electron temperature tends to be rather high (typically of the order of 1eV) to obtain an accurate representation of the Fermi function with a given number of moments, results can depart significantly from the $T = 0$ ones. This can be partially corrected for [24, 25] by using the energy functional $U - \frac{1}{2}TS$, where U is the internal energy, and S is the electron entropy. This functional approximately extrapolates the energy back to the $T = 0$ value. It is used throughout this paper for FOE. Note that this is very similar to the method of Voter *et al* [23].

The use of a finite range for the density matrix introduces errors into the forces. This can be corrected for [23], but in practice is found to be unimportant.

The utility of this approach lies in the following: the forces are easily obtained by direct differentiation of the energy since the total energy is given in terms of the global moments; the memory requirements are modest; the algorithm can be implemented very efficiently on a computer; addition of extra moments stable up to very large numbers (at least 400). These benefits, however, need to be weighed against the rather slow energy convergence with respect to number of moments. This method has been implemented on a parallel machine, and almost perfect scaling is found, as little information needs to be passed between processors.

RESULTS

Tests have been performed for energy convergence, force convergence and computational efficiency for an insulator (carbon [26] in the diamond structure), a semiconductor (silicon [27]), a metal (titanium – described by a canonical d-band model) and a molecule (benzene [28]). The full set of results is published elsewhere [29]. Here only the computational efficiency of the various methods is considered.

Figures 1 and 2 show the time per atom (per iteration for DMM) for a single total energy evaluation required to achieve a given level of accuracy for the vacancy formation energy for the four methods for carbon and titanium. It should be noted that DMM takes 2 iterations per MD step in the implementation used for this work (though for metals many more iterations may be necessary). The conditions under which the timings were taken are as follows. For DMM, the cluster size was varied. For GDOS and BOP, the calculations were performed with exact moments only. For FOE, a cluster of size four hops was taken, and the number of moments taken within the cluster was varied. Note that this will never lead to exactly converged results (a larger cluster is needed for that). However, good enough convergence for many MD simulations can be obtained in this way. For diamond, DMM is in a class all its own, being superior by far to the moments methods (see also Ref. [30]). For titanium, the exact opposite is the case. A logarithmic time scale had to be used to allow the DMM results to appear.
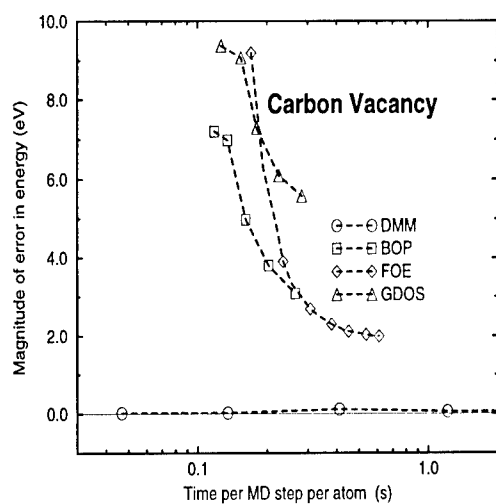
Figure 1: The error in the carbon vacancy formation energy against the time taken per MD step per atom for each of the different methods. Note the logarithmic time scale.
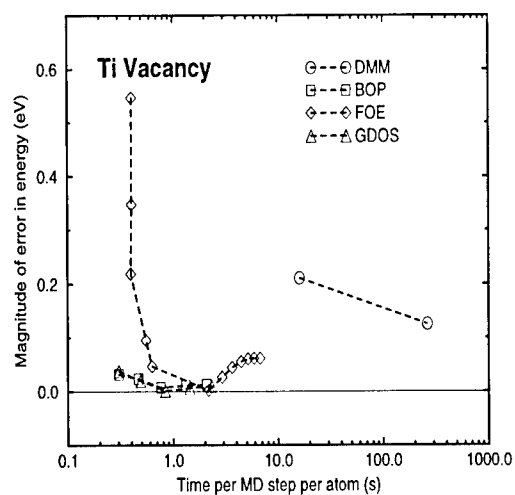
Figure 2: The error in the titanium vacancy formation energy against the time taken per MD step per atom for each of the different methods. Note the logarithmic time scale.

It is interesting to note how FOE increases its accuracy very significantly with minimal gain in time for both carbon and titanium. This follows from the simplicity of the algorithm, allowing a highly efficient implementation. It should be noted that further reduction in the times is possible by making careful use of the specific machine architecture. However, it is important to note also that a cluster of size four hops is not large enough to give the correct limiting value. This is because the addition of extra moments corresponds to a reduction in the electronic temperature, which in turn results in the increased range of the density matrix. When the density matrix extends beyond four hops, then the surface of the cluster is sampled, leading to wrong results. Increasing the cluster size will improve the accuracy, but will also increase the time of a simulation.

CONCLUSIONS

The background to the density matrix method, the bond order potential method, the global density of states method and the Fermi operator expansion have been given. There are three main points which become clear from the results of simulations: for a system involving energy gaps, DMM is the method of choice, giving excellent force and energy convergence. However, the moments methods can also give good energy convergence in reasonable times, and give sufficiently good force convergence that reliable relaxations can be performed. For liquid materials, any of the moments methods (BOP, GDOS or FOE) will give good energy convergence. However, for molecular dynamics GDOS should be used because of the high quality of its forces, and the few moments required to reach convergence. For metallic systems GDOS and BOP are suitable. FOE never gives the best performance, but also never fails. Its basic merit is that it is a reliable method that is straightforward to implement efficiently.

REFERENCES

1. X.-P. Li ,R. W. Nunes and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993)

2. M. S. Daw, Phys. Rev. B **47**,10895 (1993)

3. G. Galli and M. Parrinello, Phys. Rev. Lett. **69**, 3547 (1992)

4. F. Mauri, G. Galli and R. Car, Phys. Rev. B **47**, 9973 (1993)

5. P. Ordejon, D. Drabold, M. Grunbach and R. Martin, Phys. Rev. B **48**, 14646 (1993)

6. D. G. Pettifor, Phys. Rev. Lett. **63**, 2480 (1989)

7. M. Aoki, Phys. Rev. Lett. **71**, 3842 (1993)

8. S. Goedecker and L. Colombo, Phys. Rev. Lett. **73**, 122 (1994)

9. J. D. Kress and A. F. Voter, Phys. Rev. B **53**, 12733 (1996)

10. A. P. Horsfield, Mat. Sci. Engin. B **37**, 219 (1996)

11. E. B. Stechel, A. R. Williams and P. J. Feibelman, Phys. Rev. B **49**, 10088 (1994)

12. R. McWeeney, Rev. Mod. Phys. **32**, 335 (1960)

13. C. M. Goringe, D.Phil Thesis, Oxford University (1995)

14. J. Friedel, Adv. Phys. **3**, 446 (1954)

15. V. Heine, Sol. St. Phys. **35**, 1 (1980)

16. F Ducastelle and F. Cyrot-Lackmann, J.Phys.Chem.Solids **31**, 1295 (1970)

17. R. Haydock, Sol. St. Phys. **35**, 216 (1980)

18. S. M. Foiles, Phys. Rev. B **48**, 4287 (1993)

19. H. Hellmann, *Einfuhrung in die Quantumchemie* (Leipzig: Franz Deutsche, 1937).

20. R. P. Feynman, Phys. Rev. **56**, 340 (1939)

21. A. P. Horsfield, A. M. Bratkovsky, M. Fearn, D. G. Pettifor and M. Aoki, Phys. Rev. B **53**, 12964 (1996)

22. S. Goedecker and M. Teter, Phys. Rev. B. **51**, 9455 (1995)

23. A. F. Voter, J. D. Kress and R. N. Silver, Phys. Rev. B. **53**, 12733 (1996)

24. M. J. Gillan, J. Phys.: Condens. Matter **1**, 689 (1989)

25. A. P. Horsfield and A. M. Bratkovsky, Phys. Rev. B **53**, 15381 (1996)

26. C. H. Xu, C. Z. Wang, C. T. Chan, and K. M. Ho, J. Phys.:Condens. Matter **4**, 6047 (1992)

27. L. Goodwin, A. J. Skinner and D. G. Pettifor, Europhys. Lett. **9**, 701 (1989)

28. A .P .Horsfield, P. D. Godwin, D. G. Pettifor and A. P.Sutton, Phys. Rev. B **54**, 15773 (1996).

29. D. R. Bowler, M. Aoki, C. M. Goringe, A. P. Horsfield, D. G. Pettifor, Modelling and Simulation in Materials Science and Engineering **5**, 199 (1997).

30. J. D. Kress and A. F. Voter, Phys. Rev. B **52** 8766 (1995)

LARGE SCALE QUANTUM SIMULATIONS USING TIGHT-BINDING HAMILTONIANS AND LINEAR SCALING METHODS

GIULIA GALLI[*], JEONGNIM KIM[**], ANDREW CANNING[***] and RAINER HAERLE[*]

[*] Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), Ecublens, 1015 Lausanne, Switzerland, galli@irrma.epfl.ch
[**] Department of Physics, The Ohio State University, Columbus OH 43210.
[***] NERSC, Lawrence Berkeley Laboratory, Berkeley CA94720.

ABSTRACT

We describe linear scaling methods for electronic structure calculations and quantum molecular dynamics simulations, which are based on an orbital formulation of the electronic problem. In particular, we discuss some open problems which need to be addressed to improve the performances of these methods, and briefly review some applications to carbon and silicon systems, within a Tight-Binding framework.

INTRODUCTION

Quantum simulations are aimed at modeling materials at the microscopic level, by solving numerically the equations governing the atomic motion. In order to obtain an accurate microscopic description of most materials properties, the interaction between atoms must be described using the laws of quantum mechanics. Interatomic forces can be computed by solving the Schrödinger equation for electrons, thus determining the electronic ground state at given positions of the nuclei. In many cases of interest the nuclei can be considered as classical objects. Once atomic trajectories are determined, using, e.g., molecular dynamics, a variety of materials properties can be calculated.

The computer time required by a quantum simulation, and ultimately its feasibility, are mainly determined by the time necessary to solve the Schrödinger equation for electrons. Standard approaches[1] to the solution of this equation require a workload proportional to the cube of the number of atoms involved in the simulation: Doubling the size of the system amounts to multiplying by eight the computing time. This unfavorable scaling poses severe limitations to the kind of problems which can be tackled with quantum simulations.

Recently new methods for solving the Schrödinger equation have been developed, which imply a workload growing linearly with the system-size. These approaches, called *linear scaling methods*[2], allow one to simulate systems much larger than previously accessible, widening the range of materials science issues that can be addressed. At present, linear scaling methods using Tight-Binding Hamiltonians allow one to perform simulations involving up to thousands of atoms on small workstations, and up to ten thousand atoms for tens of picoseconds when using supercomputers. This has made it possible to study problems such as large organic molecules in water[3], thin film growth[4, 5], extended defects[6] and dislocations[7] in semiconductors. Although the implementation of first-principles linear scaling methods is less advanced than that of semi-empirical methods, promising results[8] have already appeared in the literature.

In this paper we first summarize the key features of quantum simulations based on

linear scaling approaches (section I), and then discuss technical problems involved in linear scaling electronic structure calculations, such as convergence in iterative minimizations (section II). Then we address specific issues involved in molecular dynamics simulations (section III), and give a brief review of applications within a Tight-Binding framework (section IV).

ORBITAL BASED LINEAR SCALING METHODS

So far quantum simulations based on linear scaling methods have used molecular dynamics as the methodology for the calculation of statistical properties of an ensemble of classical ions. The interaction potential between ions is then derived from the ground state of the quantum many body system of electrons, by solving the single particle Schrödinger equations, in a mean field fashion. At each MD step the forces (F_I) acting on ions are determined by computing the gradient of the total energy (E) of electrons in the field of ions, with respect to the ionic degrees of freedom:

$$F_I = -\nabla_I E[\{\mathbf{R}_I\}]. \tag{1}$$

Here $\{\mathbf{R}_I\}$ represent ionic coordinates. First-principles theories can be used to compute E, e.g. the density functional theory within some local approximation, or semiempirical single particle Hamiltonians can be adopted to compute E. Within a semiempirical Tight-Binding (TB) formulation,

$$E = E_{\text{BS}} + \sum_{LL'} V_{\text{R}}(|\mathbf{R}_L - \mathbf{R}_{L'}|). \tag{2}$$

The term V_R is a repulsive two body potential which can be derived from fits to first-principles calculations or experimental data. In standard approaches the band structure term E_{BS} is written as

$$E_{BS} = \sum_i f_i < \psi_i|\hat{H}|\psi_i >, \tag{3}$$

where $\hat{H}$ is a TB Hamiltonian and ψ_i are its eigenstates, i.e., the single particle wave functions minimizing the functional E_{BS}. The total number of electrons is N and the total number of single particle states is $N/2$, assuming double occupancy for each state ($f_i = 2$). The direct diagonalization of $\hat{H}$ needed at each step of an MD simulation requires a number of operations scaling as the third power of N, i.e. a computational workload of order N^3 ($O(N^3)$-scaling). This limits the number of atoms that can be studied in conventional TB-MD simulations to less than a few hundreds using workstations and to less than 1000, when using powerful supercomputers.

As mentioned in the introduction, in order to extend quantum MD studies to larger systems and thus to a broader class of problems, many so called $O(N)$ methods have been introduced in recent years[2], $O(N)$ meaning that their computational cost grows linearly with N and thus with the system size. Some of these approaches are based on an orbital formulation of the electronic properties[9, 10, 11] whereas others are based on the calculation of the Green's function[12, 13], the density matrix[14, 15, 16, 17] or the density[18]. In this paper we summarize the basic features of an orbital based $O(N)$ method[10]; a review of both orbital- and density matrix-based linear scaling methods is given in Ref. [2], where differences and similarities of the two approaches are discussed in detail.

A key point of O(N) methods is the evaluation of total energy and forces without computing the eigenvalues and eigenstates of $\hat{H}$. This evaluation is accomplished by dividing the full system into subsystems and then defining electronic orbitals which are *localized* in the subsystems[19]. Schrödinger-like equations are then iteratively solved for these localized degrees of freedom. The subsystems are overlapping portions of the full system, which we call localization regions. The important point is that the extension of a localization region depends on the physical and chemical properties of the system but not on the entire volume of the system. As we will see in detail in Section II, the size of a localization region is the parameter controlling the accuracy of the calculation. Using localized functions is necessary but in general not sufficient to achieve O(N) scaling. Another key ingredient is the definition of an appropriate energy functional whose minimization requires neither explicit orthonormalization of electronic orbitals, nor the inversion of an overlap matrix (**S**) between single particle wavefunctions. Such a functional is in general different from the functional E_{BS} of Eq. 2, but it has the same absolute minimum.

It has been shown in Ref. [10] that an energy functional satisfying these requirements is:

$$E_{GBS}[\{\phi\}, \mu, M] = 2 \sum_{ij=1}^{M} (2\delta_{ij} - S_{ij}) < \phi_j | \hat{H} - \mu | \phi_i > + \mu N \tag{4}$$

The matrix ($2\,\mathbf{I} - \mathbf{S}$) is the truncated series expansion of $\mathbf{S}^{-1}$ to the first order, with $S_{ij} = < \phi_i | \phi_j >$. The functional (4) depends on M orbitals - with M in general larger than the number of occupied electronic states $N/2$ -, and on a global variable μ, determining the appropriate filling of the electronic orbitals. In order to find the ground state energy, this functional is minimized with respect to electronic orbitals $\{\phi\}$, which are localized in appropriate regions of real space, the localization regions. We call these orbitals localized orbitals (LOs). By definition, the orbitals $\{\phi\}$ have finite components only inside the localization regions (LRs) and are zero outside. While the number and centers of LRs are arbitrary, typically one choses a number of LRs equal to the number of atoms, each centered at an atomic site (I). A localized orbital is then specified by its localization region, i.e. the region of space where it has finite-components, and its center, i.e. the center of the LR.

Within a TB picture, the localized orbital ϕ_i, whose center is the Ith atom, is expressed as

$$\phi_i = \sum_{J \in \{LR_I\}}^{n_b} \sum_{l} C_{Jl}^i \alpha_{Jl}, \tag{5}$$

where α_{Jl}'s are the atomic basis functions of the atom J and the index l indicates the atomic components (e.g. s, p_x, p_y or p_z for a carbon system). Here $\{LR_I\}$ indicate the set of atoms belonging to the localization region of the orbital ϕ_i.

When using localized functions, the number (n_b) of basis orbitals needed to express the electronic degrees of freedom ϕ_i becomes independent of the system size, being determined only by the extension of the localization region. Therefore the evaluation of the total energy, and of the gradients of the total energy with respect to electronic and ionic degrees of freedom, amounts to computing products of sparse matrices and thus becomes of O(N). Indeed all double sums entering, e.g. the energy expression of Eq. 4 run over orbitals defined in neighbouring localization regions. In particular, in iterative minimisations (see section II) the evaluation of the energy gradient, which in conventional iterative minimizations requires O(N^2) operations, scales as n_bO(N). Furthermore no orthogonalisation of orbitals

is needed, since one minimises a functional with implicit orthogonalisation constraints, nor inversion of overlap matrices, and therefore the whole calculation is of $O(N)$.

ELECTRONIC STRUCTURE CALCULATIONS

Total energy minimizations

The functional of Eq. 4 can be minimized using a conjugate gradient (CG) procedure, where the derivative

$$\frac{\partial E_{GBS}}{\partial \phi_i} = 4 \sum_{j}^{M} [|(H - \mu)\phi_j > (2\delta_{ji} - S_{ji}) - |\phi_j > < \phi_j|(H - \mu)|\phi_i >] \tag{6}$$

is evaluated at each iterative step. This derivative needs to be evaluated only in the localization region of ϕ_i. During the functional minimization μ is varied until the total charge of the system equals N; thus when convergence is achieved, i.e. the ground state is attained, the value of μ coincides with the electronic chemical potential. The variation of μ during the minimization procedure allows for long-range charge transfer in the minimization process, irrespective of the extent of the LRs. Therefore, the variation of μ during the minimization procedure helps to avoid being trapped at local minima, which were found in minimizations using a number of LOs equal to the number of electronic states[10]: $M = N/2$.

An example of possible traps at local minima is illustrated in Fig. 1, where we present the results[10] of a series of TB calculations using localized orbitals, for a 256 carbon atom slab. The slab, consisting of 16 layers, represents bulk diamond terminated by a C(111)-2 $\times$ 1 Pandey reconstructed surface on each side. Localization regions extend up to second neighbors and therefore contain about 20 atoms. First, conjugate gradient minimizations of the electronic structure have been performed, using two localized orbitals per LR; for carbon, having four valence electrons, this corresponds to $M = N/2$ in Eq. 4. These minimizations were carried out by starting from different wave function inputs. The only calculation which lead to a physical minimum was the one started with orbitals containing symmetry information about the system, as shown by comparing the results of Fig. 1C, left panel, with those of direct diagonalization, reported in the right panel of Fig. 1. The other calculations lead to unphysical minima: when starting with a totally random input (Figs. 1A), a local minimum with charged sites was obtained, with these sites located predominantly in the surface layers and in the middle of the slab. When starting from an atom by atom input (Fig. 1B) a local minimum was found, which corresponds to two differently charged surfaces, one positively and the other negatively charged. On the contrary, when using $M > N/2$, in particular $M = N/2 + N/4 = 3N/4$, all minimizations were found to converge to the same minimum, irrespective of the initial choice for input wavefunctions, and to yield a physical ground state charge density. This is shown in the right panel of Fig. 1, where the results obtained by direct diagonalization are compared with those of an iterative minimization with $M = 3N/4$.

Convergence of iterative minimizations

One of the major technical problems of the total energy minimizations described above is the large number of iterations needed to achieve convergence. This is substantially larger (10-100 times) than in calculations using extended states. A possible solution to the
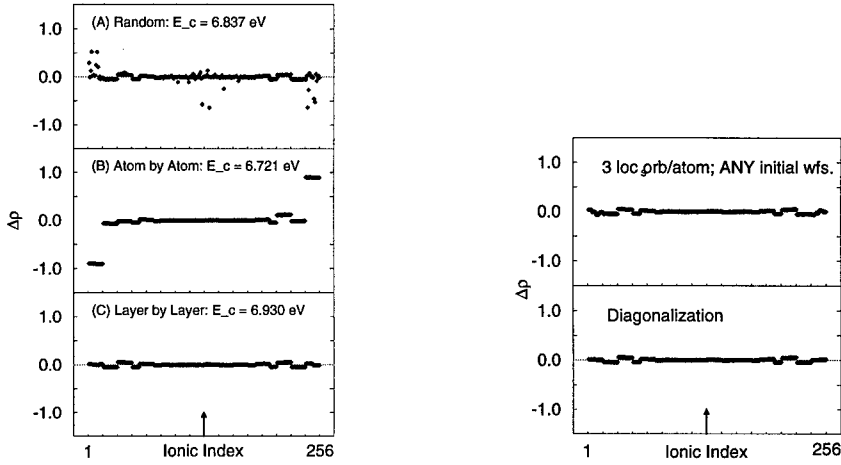
Figure 1: Differential atomic charge ($\Delta\rho$) on each atomic site of a 256 carbon atom slab. The slab, consisting of 16 layers, represents bulk diamond terminated by a C(111)-2 × 1 Pandey reconstructed surface on each side. The ionic index indicates individual atomic sites belonging to the slab, which are ordered layer by layer, starting from the uppermost surface. The arrow indicates the slab center. $\Delta\rho_K = \rho_K - \rho^0$, where $\rho_K = 2\sum_{ij=1}^{M}\sum_l <\phi_i|\alpha_{Kl}> (2\delta_{ij} - S_{ij}) <\alpha_{Kl}|\phi_j>$, $\rho^0 = 4$, and K is the atomic site (see Eq. 4-5). E_c is the cohesive energy of the slab. On the right side, in panels A, B and C we show the results of calculations performed with two localized orbitals per atomic site, i.e. $M = N/2$ in Eq. 4, and with three different wave function inputs. *Random* input: the wave function expansion coefficients (C^i_{Jl}, see Eq. (5)) on each site of a localization region (LR) are random numbers, and orbitals belonging to the same LR are orthonormalized at the beginning of the calculations. *Atom by Atom* input: each orbital has a non zero C^i_{Jl} only on the atomic site to which it is associated, and for each atomic site this coefficient is chosen to be the same. *Layer by Layer* input: each orbital has a non zero C^i_{Jl} only on the atomic site to which it is associated, and the value of this coefficient is chosen to be the same for each equivalent atom in a layer. In the case of *atom by atom* and *layer by layer* inputs, the initial wave functions are an orthonormal set. On the left side, in the upper panel we report the results of a calculation carried out with three orbitals per atomic site, i.e. $M = 3N/4 > N/2$ in Eq. 4, and with a totally random input for the initial wave functions. In this case E_c= -6.978 eV/atom. In the lower panel, we show results obtained by direct diagonalization (E_c= -7.04 eV/atom). Contrary to the calculation started from a totally random input and performed with $M = N/2$ (see right side), the calculation with $M = 3N/4$ gives a ground state charge density very close to that by diagonalization. From Ref. [10].

429

N_h	N_{LR}	E (eV/atom)
2	17	-7.186
3	41	-7.244
4	83	-7.253
$DIAG$		-7.255

Table I: Cohesive energy (E) of C in the diamond structure computed with a supercell containing 1000 atoms as a function of the number of neighbor shells (N_h) included in the definition of the localization region; the number of atoms belonging to a localization region (N_{LR}) is given in the second column. Results obtained with supercells containing 216 atoms are identical to those reported in the table. The value of E is quoted with three decimal digits to show its convergence as a function of the localization region size. The results are compared to those obtained by direct diagonalization (DIAG). Already with about 40 atoms per localization region, the error on E is less than 0.5%.

problem could come from the use of an energy functional at finite temperature[20], such as:

$$F = E_{GBS}[\phi, \mu, M] - k_B T \sum_{ij} (\mathbf{I} - \mathbf{S})_{ij} \mathbf{S}_{ij} \qquad (7)$$

where the term $k_B T \sum_{ij} (\mathbf{I} - \mathbf{S})_{ij} \mathbf{S}_{ij}$ mimics an entropy term while preserving the linear scaling of the method. Preliminary tests using the F functionals for electronic structure calculations have shown a decrease in the number of iterations needed in total energy minimizations[20]. This also implies a better conservation of energy in microcanonical molecular dynamics simulations (see section III). However the improvement is not dramatic and new schemes are needed to achieve convergence in a number of iterations comparable to those of conventional approaches.

Size of localization regions and accuracy of the calculations

The accuracy of orbital based linear scaling calculations is controlled by the size of the LR's, e.g. by their radius R_c, if LR's are chosen to be spherical. In general $\min_{\phi^{loc}} E[\phi^{loc}] = E_0^{loc} > \min_\psi E[\psi] = E_0$, but

$$\lim_{R_c \to \infty} (\min_{\phi^{loc}} E[\phi^{loc}]) = \min_\psi E[\psi]. \qquad (8)$$

By choosing appropriately the size of the localization regions, the error $\Delta E_{loc} = (E_0^{loc} - E_0)$ can be made acceptably small, as it has been shown for a variety of systems in practical calculations [9, 11, 10, 21]. Selected examples are given in Tables I-III for diamond, amorphous carbon and trans-polyacetilene, respectively.

The density matrix of systems with a gap decays exponentially with separation whereas that of metals decays with an inverse power of the separation. This was rigorously proven for one dimensional systems[22] and suggests that the error ΔE_{loc} goes to zero exponentially for insulators and algebraically for gapless systems, as the radius R_c is increased. This is

N_h	N_{LR}	E (eV/atom)
2	17	-6.241
3	34	-6.289
4	64	-6.299
DIAG		-6.307

Table II: Cohesive energy (E) of C in an amorphous carbon structure containing a mixture of sp^2 and sp^3 sites, computed with a supercell containing 512 atoms, as a function of the number of neighbor shells (N_h) included in the definition of the localization region; the number of atoms belonging to a localization region (N_{LR}) is given in the second column. The value of E is quoted with three decimal digits to show its convergence as a function of the localization region size. The results are compared to those obtained by direct diagonalization (DIAG). Similarly to diamond, when using localization regions with $\simeq$ 40 atoms, the error on E is smaller than 0.5%.
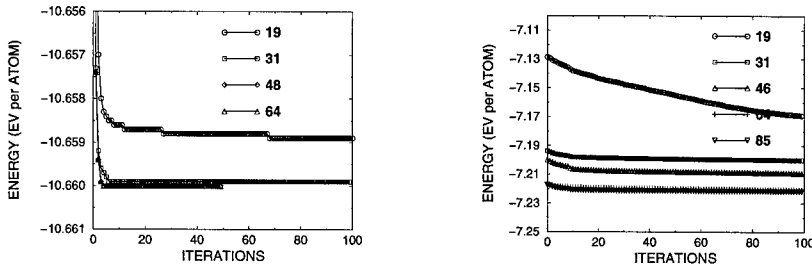
Figure 2: Convergence of the total energy as a function of the LR's size for π- and σ-2D graphite for a supercell containing 576 carbon atoms. The left hand graph shows the convergence from a random start for the σ-2D graphite for localization regions containing between 19 and 64 atoms (The value of the total energy obtained by direct diagonalization of the Hamiltonian is 10.66 eV). The right hand graph shows the convergence for π graphite for localization regions containing between 19 and 85 atoms starting from an initial state that has been converged for 100 steps with a localization region containing 19 atoms. (The value of the total energy obtained by direct diagonalization of the Hamiltonian is 7.28 eV).

N_h	E/CH unit (eV)
1	-24.679
2	-25.002
3	-25.046
4	-25.064
5	-25.073
DIAG	-25.089

Table III: Total energy per CH unit of infinite *trans*-polyacetylene (tPA) as function of the size of the localization regions, N_h, evaluated for a fixed geometry. For these calculations we set the Hubbard term of the Tight-Binding Hamiltonian to zero and used three localized orbitals per LR (i.e. $M = 3N/4$ in Eq. 4). A unit cell with periodic boundary conditions containing 40 CH units has been used. DIAG denotes results obtained by diagonalizing the Hamiltonian. From Ref.[23].

illustrated in Fig. 2 where we present the convergence of the total energy as a function of the LR's size for π- and σ-2D graphite. π-2D graphite is a metallic system, while σ-2D graphite is a model, insulating system built of C atoms with only 3 valence electrons. The profound difference in the convergence of the two systems is apparent.

MOLECULAR DYNAMICS SIMULATIONS

In MD simulations using the functional E (Eq. 4) the ground state wavefunctions $\{\phi\}$ can be obtained either by evolving the electronic states according to a Car-Parrinello [24] dynamics (see, e.g. Ref. [25]), or by minimizing the energy functional E at each ionic move. In general the atomic sites belonging to a given localization region vary as a function of time, when, e.g. atoms diffuse or change their local coordination. This implies an abrupt modification of the basis functions used for the expansion of the LO's $\{\phi\}$ and therefore a discontinuity of $\{\phi\}$ as a function of the ionic positions. When the set of atoms belonging to a LR_I change, E must be minimized with respect to the electronic degrees of freedom. Therefore, in most cases it is convenient to minimize the energy functional at each ionic step, irrespective of whether the LR changes at that step, instead of using a Car-Parrinello dynamics.

Within a Tight-Binding picture, when performing MD simulations particular care must be taken in preventing unphysical charge transfers between neighboring atoms. An approximate way to control charge transfers consists in adding to the Hamiltonian $\hat{H}$ an Hubbard-like term [26] $H_U = U \sum_I (q_I - q_I^0)^2$, where q_I is the Mulliken population at atomic site I, q_I^0 equals the valence atomic charge and U is a constant chosen on a semiempirical basis. The Mulliken population is given by $q_I = 2 \sum_{ij} (2\delta_{ij} - S_{ij}) < \phi_i | R_I > < R_I | \phi_j >$, where $< \phi_i | R_I >$ indicates the projection of the localized orbital ϕ_i onto the localization region associated to atom I. In the presence of a Hubbard like term, the line minimization required in a CG procedure amounts to the minimization of a polynomial of eight degree in the variation of the wave function along the conjugate direction. In the absence of the term H_U the polynomial is of third degree. The line minimization can be performed by

432
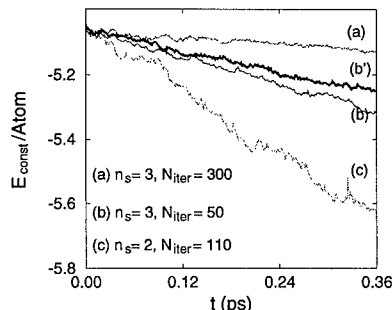
Figure 3: Energy per atom (E_{const}) as a function of the simulation time (t) for constant temperature molecular dynamics (MD) simulations of liquid C. The simulations was performed using a Nosé thermostat. We used a supercell containing 64 atoms. $E_{const}=E_{kin} + E_{GBS} + E_{thrms}$, where E_{kin} is the ionic kinetic energy, E_{GBS} is given by Eq. (4) and E_{thrms} is the sum of the potential and kinetic energies associated to the Nosé thermostat. The LRs contain about 20 atoms. Lines (a) and (b) correspond to MD runs with $M = 3N/4$ ($M > N/2$, $n_s = 3$) whereas line (c) corresponds to a simulation with $M = N/2$ ($n_s = 2$). Line (b') corresponds to the same simulation as line (b) but performed using the functional F (see Eq. 7), with $K_B T = 1$ eV. The time step used in the three MD runs was 30 a.u.(0.73 fs). At each step, the electronic structure was minimized by a conjugate gradient procedure with a fixed number of iterations (N_{iter}). The simulations represented by lines (b) and (c) require the same computational cost.

evaluating the coefficients of the polynomial. We note that the O(N)-TB method is particularly well suited to MD simulations in which the number of particles changes during the run. An extra-particle can be added to a N_c-particles system, at a distance where it does not interact with any of the N_c particles, because of the localization constraints imposed on the electronic orbitals. Thus the set of LO's minimizing the total energy of the ($N_c + 1$)-particle system is initially given by the union of the sets of LOs $\{\phi\}$'s minimizing respectively the E_{GBS} of the N_c and that of the extra-particle systems. A subsequent gradual approach of an extra-particle to the N_c particle system only weakly perturbs its electronic ground state. Therefore the minimization of the TB-Hamiltonian for the ($N_c + 1$)-particle system takes only a small number of steps, i.e. much smaller than any optimization from scratch of the energy of the ($N_c + 1$)-particle system. This has been used, e.g., in Ref. [5] to study the growth of a small fullerene film in the gas phase and in Ref. [4] to study the growth of a thin film on a surface.

In general, in MD simulations using linear scaling methods conserved quantities (such as the total energy in a microcanonical simulation) are obtained to a lesser degree of accuracy, compared to simulations using diagonalization procedures for the electronic ground state. This reflects the problem of the large number of iterations needed to minimize E_{GBS} when using localized orbitals (section II). We show an example in Fig. 3, which displays the energy per atom as a function of simulation time for constant temperature MD of liquid carbon[10], which is a disordered metallic system, i.e. a system rather difficult to treat within a O(N) scheme. The conservation of energy is satisfactory when using a reasonable number of iterations (i.e. 50 iterations/steps) but not as good as the one usually obtained
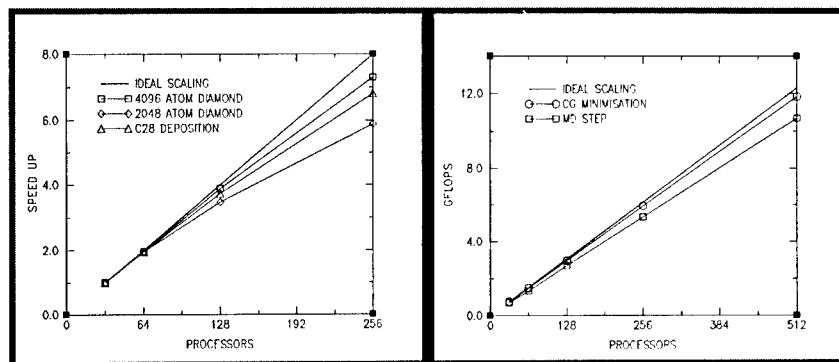
Figure 4: Right panel: 'Weak scaling' of the O(N)-Tight-Binding Molecular Dynamics(MD) code implemented in Ref. [28] for a bulk diamond system of 2048 to 32768 atoms running on 32 to 512 processors of a cray T3D. The localization region for each orbital contains 45 sites. The curce of circles and squares show the performance for the conjugate gradient (CG) minimization and for a full MD run, respectively. The ideal scaling is shown for the CG curve. Left panel: 'Strong scaling' of the same code for two bulk diamond systems of 2048 and 4096 atoms, and an inhomogeneous system of fullerenes deposited on a diamond surface (see Fig. 5). All orbital localization regions are taken up to third nearest neighbors. Similar speed up curves were found for the Cray T3E900 with the overall speed being 3.2 to 3.5 times faster than the T3D, depending on the system size.

with diagonalization techniques. The picture also shows the improvement obtained when using a finite temperature energy functional of Eq. (7).

RECENT APPLICATIONS

So far, most large-scale applications with $O(N)$ methods have been carried out using TB Hamiltonians. Molecular dynamics simulations involving about one thousand atoms can routinely be performed on workstations [5, 27], while the efficient implementations of the O(N)-TB algorithm on a massively parallel computer[28, 29] allows one to simulate 5000-10000 atoms for tens of picoseconds[4]. The performance of a parallel code[28] implemented on a T3D/E is shown in Fig. 4.

Here we briefly review applications of linear scaling methods to the study of carbon and silicon systems, in particular carbon clusters on surfaces and defects in silicon.

Regarding clusters on surfaces, two simulations have appeared in the literature: The simulation of C_{60} impacts on a semiconducting substrate[27], involving more than 1000 atoms and carried out on a small workstation, and the study of C_{28}'s thin film deposition on a non-metallic surface, involving about 5000 atoms and carried out on a parallel supercomputer[4].

In Ref.[27], Tight-Binding molecular dynamics simulations of C_{60} impacts on the reconstructed diamond(111) surface were presented, which provided a detailed characterization of the microscopic processes occuring during the collision, and allowed one to identify three impact regimes: A regime in which the molecule bounces off the surface (for incident ki-
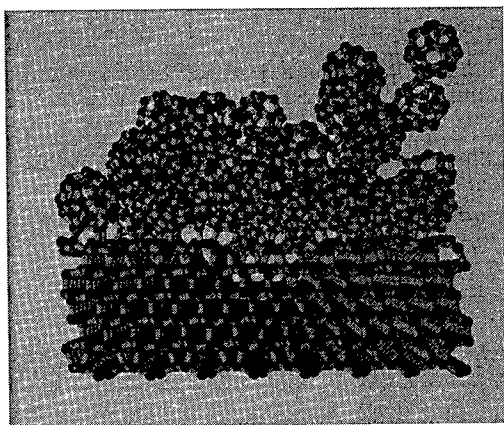
Figure 5: A snapshot of the full system (4472 atoms) at the end of a molecular dynamics simulation of the deposition of C_{28} cages on a diamond substrate, showing the undamaged clusters. From Ref. [4].
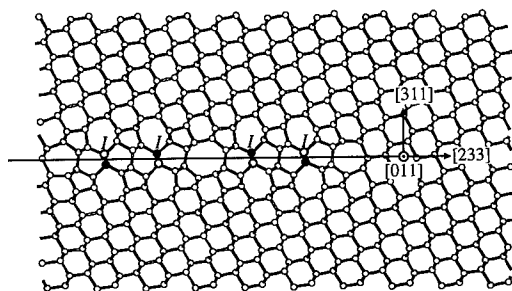
Figure 6: Projection on the $\{0\bar{1}1\}$ plane of an extended $\{311\}$ defect in Silicon. This $\{311\}$ defect is formed by condensation of the interstitial chains along the [011] direction (I) side by side in the $\langle 233 \rangle$ direction. In Ref. [6] it was shown that the defect structure is energetically more favorable than isolated interstitials due to: (i) very low dangling bond concentration and (ii) small bond length and bond angle distortions by introducing five- and seven-member rings upon the insertion of the interstitial chains. From Ref. [6].

netic energies, E_{in}, smaller than 120 eV), another one in which the molecule can be either adsorbed on or leave the substrate, being undamaged, and a third regime (E_{in} 240 eV where the molecule breaks up after the impact and pieces of the broken cage form stable bonds with the substrate.

In Ref. [4], it has been shown that C_{28} fullerenes can be deposited on a surface to form a thin film (see Fig. 5) of nearly defect free molecules, which act as carbon superatoms. These findings help clarify the structure of disordered small fullerene films and also support the recently proposed hyperdiamond model for solid C_{28}. Disordered and ordered solid forms of C_{28}'s have been studied in detail in Ref. [5].

Recently the O(N)-method has also been used to study defects[6] and dislocations[7] in Si. An example of an extended defect studied in Ref. [6] is presented in Fig. 6.

CONCLUSIONS

The linear scaling methods described here are a promising framework for the study of broader classes of problems than is affordable with conventional techniques. The implementation of $O(N)$ methods for tight-binding and semiempirical Hamiltonians has already given remarkable results, including the study of thin films[4, 5] and large organic molecules in water[3].

One of the major technical problems of $O(N)$ calculations based on localized functions is the large number of iterations needed to converge the electronic structure at each molecular dynamics step. This is significantly larger than in calculations using extended states[10]. A possible solution to the problem could come from the use of finite temperature calculations. A more fundamental problem concerns the computation of spectral properties, since eigenstates and eigenfunctions are not computed when using $O(N)$ methods. This might be accomplished by combining the present $O(N)$ methods with inverse iteration approaches[30].

As a final remark, we note that as the size of the system which can be simulated increases, the time scales involved in the problem can change considerably and become unaffordable using conventional simulations schemes. Therefore the development of approaches extending the application of quantum MD to larger systems must be accompanied by the developments of methods capable of handling slower time scales, in order to build efficient simulation tools.

ACNKOWLEDGEMENTS

REFERENCES

1. For a review see, e.g., G. Galli and A. Pasquarello, in Computer Simulation in Chemical Physics, Edited by M.P.Allen and D.J.Tildesley, p. 261, Kluwer, Dordrecht (1993), and M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1993).

2. For a review see, e.g., G. Galli, Current Opinion in Solid State and Material Science vol.1 **6**, 864 (1996).

3. T.S.Lee, D.M.York and W. Yang, J. Chem. Phys. **105**, 2744 (1996).

4. A. Canning, G. Galli and J. Kim, Phys. Rev. Lett. **78**, 4442 (1997).

5. J. Kim, G. Galli, J. Wilkins and A. Canning, J. Chem. Phys. 1997 (in press).

6. J. Kim, J. W. Wilkins, F. S. Khan and A. Canning, Phys. Rev. B **55**, 16186 (1997).

7. J. Bennetto, R. W.Nunes and D. Vanderbilt, Phys. Rev. Lett. **79**, 245 (1997).

8. P. Ordejón, E. Artacho and J. M. Soler, Phys. Rev. B **53**, R10441 (1996).

9. F. Mauri, G. Galli Phys. Rev. B **50**, 4316 (1994).

10. J. Kim, F. Mauri, G. Galli, Phys. Rev. B **52**, 1640 (1995).

11. P. Ordejón, D. Drabold, R. Martin and M. Grunbach, Phys. Rev. B **51**, 1456 (1995).

12. S. Baroni and P. Giannozzi, Europhys. Lett. **17**, 547 (1991).

13. M. Aoki, Phys. Rev. Lett. **71**, 3842 (1993).

14. X.-P. Li, R. Nunes, and D. Vanderbilt, Phys. Rev. B **47**, 10891 (1993); M. S. Daw, Phys. Rev. B **47**, 10895 (1993).

15. S. Goedecker, and L. Colombo, Phys. Rev. Lett. 73, 122 (1994).

16. E. Hernandez, M.J. Gillan, and C.M. Goringe, Physical Review B **53**, 7147 (1996).

17. W. Kohn, Phys. Rev. Lett.**76**, 3168 (1996).

18. M. Foley and P. A. Madden Phys. Rev. B **53**, 10589 (1996) and references therein.

19. G. Galli and M.Parrinello, Phys. Rev. Lett. **69**, 3547 (1992).

20. J. Kim, F. Mauri and G. Galli, unpublished.

21. S.-Y. Qiu, C.Z. Wang, and C.T. Chan, Journal of Physics. Condensed Matter **6**, 9153 (1994).

22. W. Kohn, Phys. Rev. **115**, 809 (1959).

23. M. Elstner, G. Galli and S. Suhai, unpublished; M. Elstner, S. Suhai and G. Galli, Physica D 1997 (in press).

24. R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

25. F. Mauri, G. Galli, and R. Car, Phys. Rev. B **47**, 9973 (1993).

26. C. Xu et al., J. Phys. Condensed Matter **4**, 6047 (1992).

27. (a) G. Galli and F. Mauri, Phys. Rev. Lett. **73**, 3471 (1994); (b) G. Galli, A. Canning and F. Mauri, in "Materials Modelling and Design", Eds. Vijay Kumar, Surajit Sengupta and Baldev Raj, "Springer Proceedings in Physics" (Springer Verlag, Heidelberg, 1997).

28. A. Canning, G. Galli, F. Mauri, A. De Vita and R. Car, Comp. Phys. Comm. **94**, 89 (1996).

29. S. Ito, P. Ordejón, R. M. Martin, Comp. Phys. Comm. **88**, 173 (1995).

30. K. A. Mäder, L.-W. Wang and A. Zunger, Phys. Rev. Lett. **74**, 2555 (1995) and references therein.

TIGHT BINDING SIMULATIONS OF DISORDERED SYSTEMS

V. ROSATO, M. CELINO
ENEA, "Casaccia" Research Centre, High Performance Computing and Networking Project,
SP59 P.O.Box 2400, 00100 ROMA A.D. (Italy)

ABSTRACT

Carbon allotropy allows the formation of a large variety of disordered structures whose properties depend on the density and on the preparation technique. Computer simulations, based on quantum mechanics, can reproduce the essential features of these systems. We first assess the reliability of the amorphous carbon structures generated by molecular dynamics simulations based on a semi-empirical Tight Binding hamiltonian. Then we attempt to understand of the path followed by diamond to reach the amorphous phase via a direct crystalline-to-amorphous transformation. Results concern the large variety of structures which can be obtained and a general adequacy of the semi-empirical Tight Binding hamiltonian to reproduce the essential features of amorphous carbon structure. Moreover, it is shown that the process of direct transformation from the crystalline into the amorphous phase occurs continuously. The formation of three-fold coordinated sites is not followed by an immediate site re-hybridization. When this process takes place, a large strain sites associated with the mis-coordinated sites is released.

INTRODUCTION

Allotropy is at the origin of the intriguing topology of the disordered forms of carbon. The reproduction of the various chemical bonds in amorphous carbon (a-C hereafter) is a challenge for simulation studies based on microscopic models [1]-[8]. The complex nature of the carbon chemical bond makes simulation models based on empirical interatomic potentials [9] inadequate to reproduce the basic properties of the a-C structures (radial distribution function, bond distribution, coordination). A better agreement with experimental results has been achieved, in turn, by the Car-Parrinello (CP) method [1, 6, 7]. Recent studies [10] have pointed out the good qualitative agreement between the experimental electronic density of states (EDOS) and the corresponding functions evaluated on simulated systems using the CP approach. However, the study of disordered structures compels the use of large model systems ($N \sim 10^3$). Such large size systems are computationally too expensive for ab-initio techniques. In this respect, it is important to assess the reliability of amorphous structures produced by approximate quantum-mechanical model hamiltonians, such as semi-empirical [2, 4, 8] and ab-initio Tight Binding (TB) [5] and to deploy them for more systematic studies.

The need for systematic studies is related to the large variety of structures and properties which have been measured on a-C samples at different densities [11, 12, 13]. It is, thus, important that simulation methods produce manageable representations of the amorphous structures in the whole density range which favourably compare to experimental results.

Among the problems which simulation studies must face is the preparation of the amorphous structure. This task is usually performed by heating a crystalline solid above the melting point and by subsequently quenching it down to room temperature with a given quench rate. This technique has, however, several drawbacks: (a) it is very demanding from the computational point of view, in particular when adopted in conjunction with sophisticated model hamiltonians; (b) the structure resulting from this process depends on the quench rate [8].

The two major goals of this work will be

1. assessing the limits of usability of Tight Binding to produce a-C structures which could be compared to ab-initio results, to review the attempts made so far on this topic and to discuss the obtained results;

2. trying to address a more fundamental question concerning the way a diamond crystalline structure reaches its amorphous phase. Given the cited proprierty of allotropy, it is relevant to understand the path followed by a system, which has many different ways to accomodate the presence of disorder, when undergoing a transition from a crystalline to a fully disordered phase.

We will then study the amorphization process during a *direct* transformation from the crystalline to the amorphous phase. The transition is triggered by inserting an increasing number of dissociated Frenkel pairs (i.e. an interstitial-vacancy pair) into a diamond lattice. This technique has the advantage of allowing a constant monitoring of the response of the crystalline phase to the presence of structural disorder and to correlate the occurrence of the transition to the critical behavior of some physical quantity. This approach should answer several questions:

1. to what extent carbon allotropy allows the system to retain more symmetric crystalline structures with respect to transformation into a disordered structure?

2. which is the role played by the strain imposed to both the ionic and the electronic structure?

3. does a continuous re-hybridization of low coordination sites occur as disorder increases or the formation on sp^2 sites takes place at once at the crystalline-to-amorphous transition?

The scheme of the paper is as follows: in the next section, the characteristics of the amorphous structures resulting from different simulation techniques, which make use of a semi-empirical Tight Binding hamiltonian [14, 15] (quench of a melt, combination of Reverse Monte Carlo and simulated annealing, direct amorphization from the crystalline phase via introduction of Frenkel pairs) will be shown. Results will be discussed and compared to corresponding structures generated via other model Hamiltonians (in particular via CP technique). In the subsequent section, the path followed by the crystalline structure to reach the limit of instability, in the case of Frenkel pairs insertion, will be shown and the corresponding behavior of relevant quantities (i.e. local pressure) will be correlated to the onset of the transformation. The presence and the characteristics of the electronis states which appear in the gap as a consequence of the onset of disorder will be also investigated. The properties of gap states will be related to the structures of the disordered regions where such states are localized.

All the computations reported in this study were performed by semi-empirical Tight Binding Molecular Dynamics (TBMD) simulations [15] using the TB parametrization by Kwon et al. [16].

AMORPHIZATION TECHNIQUES

In most simulation studies amorphous structures are generated by rapidly quenching the system from the molten state to room temperature. The quench rates are slow enough to allow the system to undergo the liquid-to-amorphous transition. However, the computational workload required to firstly produce a molten phase and, then, to perform the subsequent quench of the liquid rapidly increases with the system size and the complexity of the Hamiltonian. Less common techniques to prepare amorphous systems are based on the introduction of a non-equilibrium concentration of interstitial atoms into crystalline structures [17, 18] or antisite defects in binary or multi-atom systems [19], whose presence

440

destabilizes the structure thus triggering the onset of a crystalline-to-amorphous transition [20]. The Reverse MonteCarlo technique [21] has been also used to produce a-C structures [22, 23]. In the following, we will discuss the results of three different amorphization techniques used to produce a-C structures.

Quench of a melt

A diamond structure, containing $N = 216$ atoms, arbitrarily set at a density of 3 gr·cm^{-3} was firstly heated to $T = 5000$ K at fixed volume. At this temperature and density the crystalline solid melts. The system was subsequently equilibrated at fixed volume and temperature for a few thousands time steps. The system was then quenched down to room temperature with different quench rates. The results reported in table 1 show that the amorphous structures reached by fast quench rates are characterized by a larger fraction of sp^3 sites with respect to those obtained by slow quench rates [8]. Moreover, it can be be noted that the system undergoes a complete re-hybridization for quench rates as low as $q \sim 10^{14}$ K sec^{-1}. This leads to a complete sp^2 bonded structure forming a disordered graphitic-like network [8]. The tendency of the system to assume an increasing three-fold coordination with slow quench rates was discussed also in [7]. A direct comparison between TB and CP data put in evidence the general tendency of TB to underestimate the fraction of sp^3 sites which form in the structure at the given density, independently on the quench rate. Despite this partial disagreement, the EDOS resulting from TB can be favourably compared to that obtained by CP technique, as far as the position of the main peaks is concerned. In particular, it has been shown the origin of the small peak appearing at the upper edge of the valence band related to the presence of π-electrons [7, 10, 24].

Reverse Monte Carlo

Several authors have noted the ability of the Reverse Monte Carlo (RMC) technique [21, 22] to produce 3-dimensional (3D) structural models of disordered materials which agree with available diffraction data. In essence, the RMC algorithm is an optimization technique which consists of displacing point particles in a box of given dimensions (fixed by the requirement of the overall number density) until the derived radial distribution function $G(r)$ matches the experimentally measured data. This technique is totally divorced from interaction schemes and allows complete data sets from different sources (neutron or X-ray diffraction, chemical constraints) to be fitted. In the RMC technique, starting from an initial guess configuration to which is associated the model radial distribution function $G_{mod}(r)$, a "quality of hit" value Π^2, defined by the function

$$\Pi^2 = \sum_{i=1,n} \alpha_i [G_{mod}(r_i) - G_{exp}(r_i)]^2 \qquad (1)$$

can be associated to the guess structure. In eq.(1) n is the number of data points in the experimental radial distribution funtion $G_{exp}(r)$ and α_i suitable weighting factors to help convergence in the different regions of the distribution.

The RMC technique consists in randomly displacing the positions of the atoms with the constraint of minimizing the value of Π^2. The quality of hit can be then modified by introducing, into the optimization procedure, further requirements inferred from experimental data (e.g. bond angles, bond lenghts etc.). Recent applications of this technique to the determination of the tetrahedral a-C structure have introduced further geometrical constraints, such as the presence of different types of C atoms in the structure related to the different bonding configurations. [22, 23]. The simultaneous fulfillment of many different constraints largely increases the computational workload of the algorithm but allows a rather good agreement with the available experimental data. In this form, the method entirely relies on the availability of different experimental data.
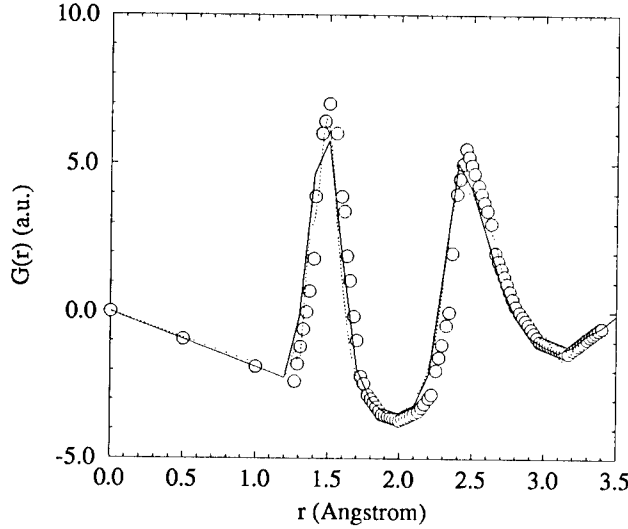
Figure 1: Experimental radial distribution function $G(r)$ (circles) [13], compared to that of an as-prepared structure from RMC (full line) and to that resulting from the subsequent structure relaxation performed via TBMD (dotted line).

In a recent work [24], in turn, a different strategy was adopted. The RMC algorithm is applied, in that case, in its simplest form, by using only a limited set of experimental data. Other than the radial distribution function, the only further constraint was to produce a structure with a desired ratio of sp^2 and sp^3 sites. The aim was to use RMC to produce starting configurations to be subsequently used in MD simulations based on a accurate model Hamiltonian scheme [16]. This strategy has, at least, three advantages:

1. it relies on very reliable data (radial distribution function) and makes no further assumption on the topology of the structure;

2. coupled with an accurate model Hamiltonian, it can be predictive, as it produces a set of structural data which can help the interpretation of further experimental data (e.g. electronic structure) not included in the optimization strategy;

3. it considerably reduces the computational time required to produce an amorphous structure (with respect to that needed with usual techniques, e.g. quench-from-a-melt).

This implementation of the RMC technique was further optimized by using a simulated annealing strategy to find the global minimum of the function Π^2 of Eq.1 [24].

At the end of the RMC procedure, the system is found within a given basin of solutions, which is very similar to that reached by the quench-of-a-melt technique [8, 24].

The system was subsequently equilibrated, with TBMD, at room temperature. Figure 1 shows a typical $G_{mod}(r)$ of a structure resulting from the RMC optimization process (full line) and that which results after a subsequent equilibration period (dotted line). These curves are compared with the experimental curve [13]. Results in terms of potential energy, bond angle distribution, equilibrium fraction of sp^2 and sp^3 sites agree with the corresponding values calculated on a-C structures obtained by quenching the melt with $q \sim 10^{15}$ K sec^{-1}.

TB dynamics		quench-rate (K/sec)	U_q (eV/atom)	U_r (eV/atom)	n_2	n_3	n_4
	from ab-initio MD liquid	q1=8.8 10^16	-7.470	-7.639	1.6	69.0	29.3
		q6=2.6 10^16	-7.598	-7.602	4.6	50.7	44.6
		q2=8.8 10^15	-7.629	-7.636	4.7	50.0	46.3
		q3=1.7 10^15	-7.623	-7.628	4.7	63.5	32.8
		q4=4.4 10^14	-7.907	-7.915		100.0	
	from TBMD liquid	q1=8.8 10^16	-7.418	-7.622	5.2	70.4	23.6
		q2=8.8 10^15	-7.683	-7.699	2.8	71.3	25.9
		q3=1.7 10^15	-7.795	-7.794	1.4	70.4	28.2
		q4=4.4 10^14	-7.877	-7.876	3.7	94.4	1.8
CP dynamics Ref.[7]	from ab-initio MD liquid	q5= 6.5 10^19			-	35.0	65.0
		q6= 2.6 10^16			-	34.0	66.0
		q7=9.4 10^15			-	40.0	58.0

Table I: Relevant thermodynamic and structural data for amorphous structures obtained by the quench of a melt technique. The rows comprised under "TB dynamics" contain the results of the quench of liquid structures obtained via a TBMD formalism, starting from ab-initio (CP) liquid [6] (shaded region) and from TB liquid (lower region). The rows comprised under "CP dynamics" simply report data collected from Refs.[6, 7]. U_q and U_r are the energy per particle of the as-quenched and "aged" amorphous structures, respectively. The n_α values (in percent) are the fraction of α-coordinated sites after the $T = 300K$ relaxation process.

Introduction of Frenkel pairs

A diamond structure containing $N = 216$ atoms has been firstly equilibrated at $T = 300K$ at fixed temperature and vanishing external pressure. Starting from the resulting configuration, several structures, containing an increasing concentration c of dissociated Frenkel pairs (c = 0.5%, 1%, 3%, 5%, 8%, 12%, 15%, 18%) have been produced. Dissociation have been imposed by avoiding the insertion of the interstitial in nearest neighbour position of some vacancy (to inhibit direct recombination).

We have then equilibrated at constant temperature and pressure the disordered structures during nearly 10,000 time steps. At the end of the equilibration period, the system has been monitored for a few thousands time steps. The results refer to the average values of thermodynamic, structural and electronic properties during that period.

From the thermodynamic point of view, the simulation cell undergo an increasingly large distortion. This effect induces a potential energy increase (fig.2) and a continuous density decrease (fig.3). Both these quantities saturate at large c to values typical of high density a-C structures (see, e.g. table 1 and ref.[23]), although a fully-developped a-C phase is achieved only for c $\sim$ 0.15. Below this value, indeed, the structure retains its crystalline character, although presenting a diffuse disorder. The potential energy per atom at c=0.18 has values in the same range as those resulting from other amorphization procedures (see, e.g. Table 1). At that c value, also the bond angle distribution (fig.4) is quite similar to those reported in [6, 24]. Aside from the information on the thermodynamic behavior, MD simulations allow to study structural properties of the system during the disordering process. Figure 5 reports the pair correlation functions of the systems at c=0, 0.08 and 0.18. At c=0.08 the structure is still crystalline, although presenting a diffuse disorder. In the case of c=0.18, in turn, the structure is amorphous.
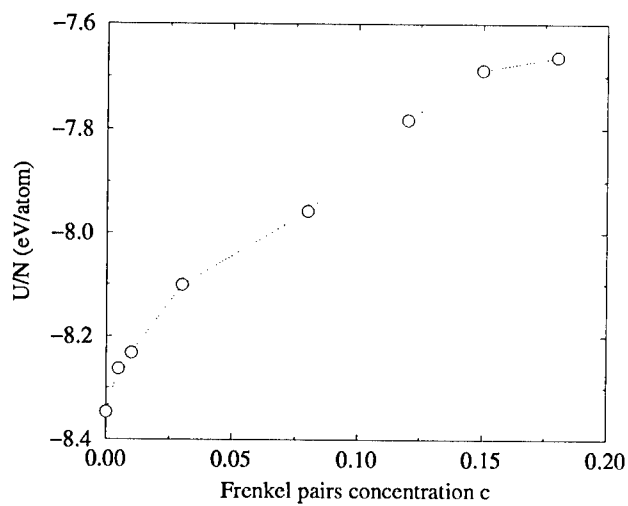
Figure 2: Potential energy per atom of the carbon system as a function of the concentration c of Frenkel pairs introduced into the structure. All simulations have been performed at $T = 300K$ and vanishing external pressure.
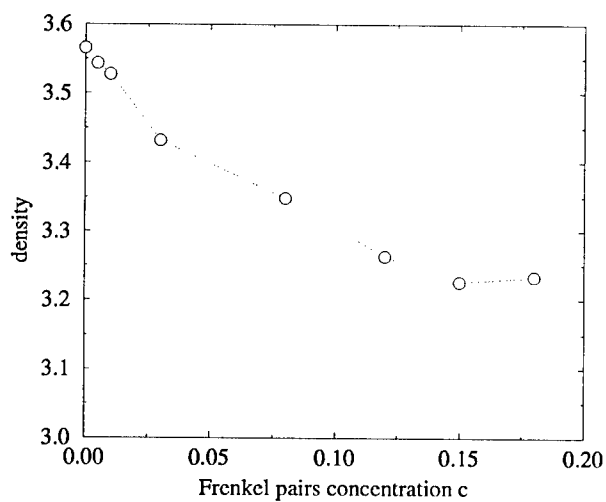
Figure 3: Variation of the density δ (expressed in gr cm^{-3}) of the disordered phase as a function of the concentration of Frenkel pairs introduced into the structure. All simulations have been performed at $T = 300K$ and vanishing external pressure.
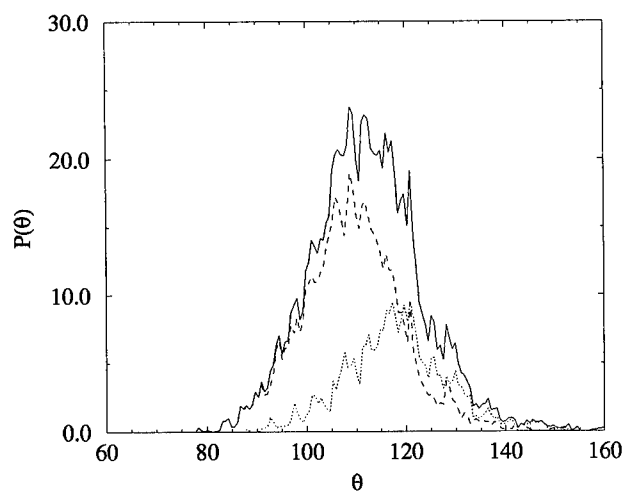
Figure 4: Bond angle distribution P(θ) of the relaxed **c**=0.18 configuration. Full line= total distribution, dotted line=bond angle distribution of three-fold sites; dashed line= bond angle distribution of four-fold sites.
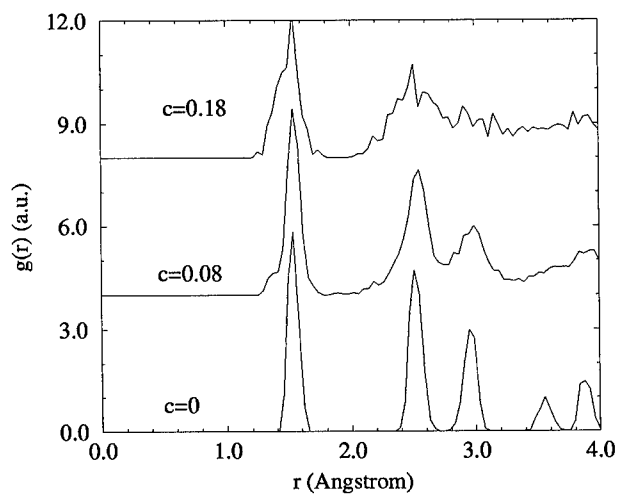
Figure 5: Pair correlation functions of diamond during the disordering process, at **c**=0 (perfect crystal at T=300K), **c**=0.08 and **c**=0.18 (above the amorphization threshold).
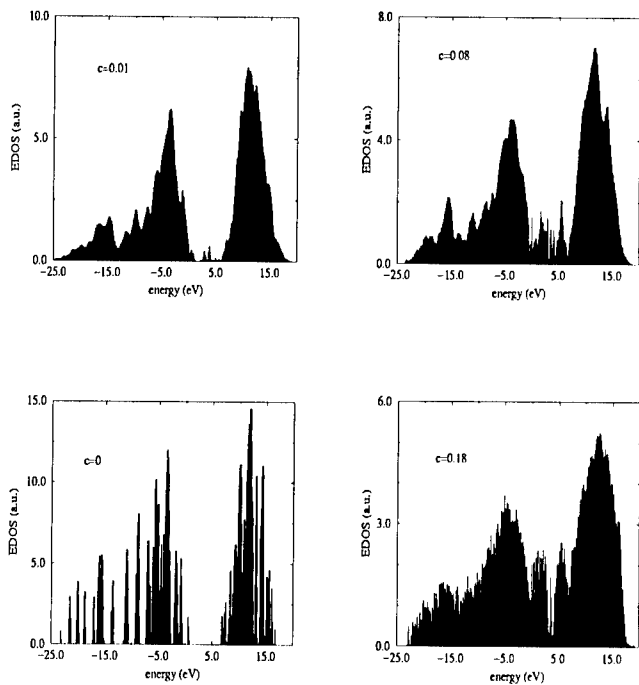
Figure 6: EDOS evaluated at equilibrium on systems at different values of c (c=0, 0.01, 0.08, 0.18).

AMORPHIZATION PATH

The electronic structure of the disordered systems shows the presence of a sparse set of electronic states within the band gap. Figure 6 shows the EDOS of disordered systems at different values of c. The edges of both the valence and conduction band are hidden by both the band broadening and by this set of states. In order to evaluate the band gap width as a function of c, the Tauc construction [25] has been used in the cases where a large amount of disorder was already established. All the structures, at different values of c are characterized by a concentration of the three-fold coordinated sites proportional to the disorder parameter c. The width of the band gap and of the homo-lumo gap, as a function of c, has been reported in fig.7. The fraction of three-fold coordinated sites is also reported in fig.7. These data confirm the empirical rule that the optical (Tauc) gap depends linearly on the sp^2 fraction [26]. The homo-lumo gap appears to drop to small values (~ 0.5 eV) since the beginning of the disordering process.

The gap states, recently studied also by Robertson [26, 27], are characterized by a large value of the participation ratio λ_i (expressed as $\lambda_i = \sum_{j=1,n} t_{ji}^4$ where t_{ji} are the elements of the eigenvector matrix) of the corresponding eigenvalues. These states are, in fact, localized around the atomic sites with coordination different than four-fold. In fig.8 two different plots relative to the case c=0.05 are shown: the lower one represents the EDOS region containing the gap; the upper one reports the localization value λ_i of the corresponding
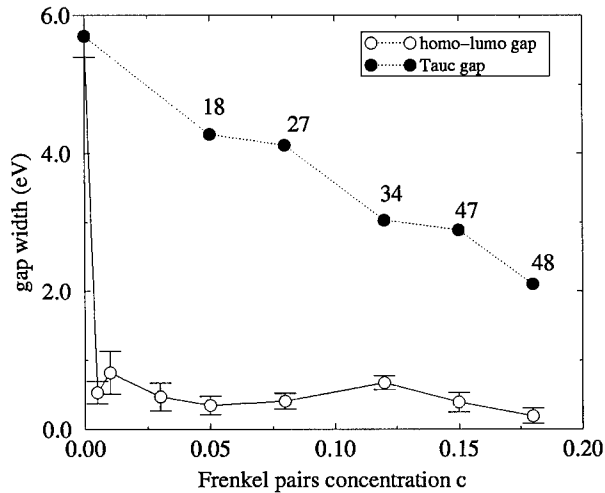
Figure 7: Behavior of the width homo-lumo gap (empty circles) and Tauc (optical) gap (full circles) as a function of **c**. The number appearing near the full circles represents the concentration of the three-fold sites in the system.

eigenvalues. The larger values of λ_i express the localization of the eigenvalues localized on three-fold coordinated sites.

As noted above, the system reacts to the insertion of Frenkel pairs by creating a number of three-fold coordinated sites (nearby the introduced vacancies). The formation energy of the dissociated Frenkel pair resulted to be E_F =17.3 eV, in agreement with previous ab-initio calculations [28]. The structure relaxes around the introduced defects (vacancies and interstitials) so as to mantain the four-fold coordination throught the rest of the system. However, the appearence of three-fold coordination sites is not related, at least in the early stages of the disordering process, to a site re-hybridization. This fact can be put in evidence by looking at the bond-angle distribution $P(\theta)$ of the three-fold coordinated sites. In fig.9 the $P(\theta)$ of the three-fold sites for **c** $= 0.01$ is shown. A peak at $\theta \sim 109°$ (typical of sp^3 sites) dominates the distribution, pointing out that the system is still able to keep the sp^3 hybridization by creating dangling bonds and by partially distorting the hybrid structure. The distortion is also indicated by the minor peak at $\theta \sim 145°$ in the distribution, related to distorted sp^3 hybrids which tends to assume a nearly planar form typical of sp^2 bonds with asymmetric lobes at $\theta \sim 104°$ and $\theta \sim 150°$.

A useful measure of the local state of hydrostatic stress around the mis-coordinated atoms is represented by the so called "local pressure", derived from the trace of the atomic-level stress tensor [29]. The effect of the lattice distortion associated with three-fold coordinated atoms on the overall stability of the crystalline structure is shown in fig.10, where the local pressures are displayed as a function of the site number. The mis-coordinated sites which could not perform a complete re-hybridization show anomalous levels of local pressure with respect to undistorted, four-fold coordinated sites. As soon as re-hybridization occurs, the local pressure level associated with mis-coordinated sites decreases, as the electronic-structure rearrangement releases the strain.

As soon as the degree of disorder increases, above **c** $= 0.03$, the $\theta = 109°$ and $\theta = 145°$ peaks disappear and the dominant peak becomes $\theta = 120°$. This transition testifies the occurrence of a complete re-hybridization of the three-fold sites and the appearence of "true" sp^2 hybrids. At a critical level of disorder, estimated to be at **c** ~ 0.15, the struc-

447
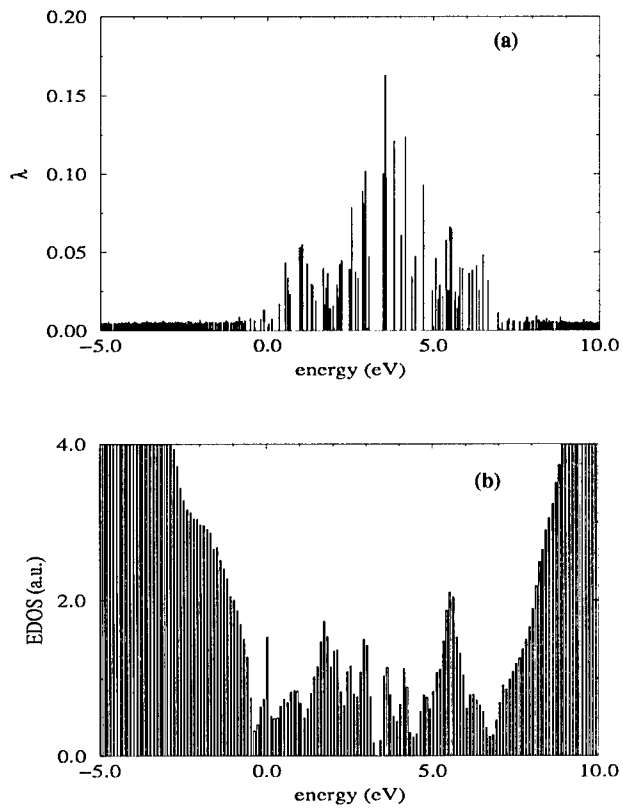
Figure 8: Panel (a): participation ratio of the different eigenvalues in the region of the gap. Panel (b) shows the corresponding part of the EDOS. Both the results refer to the case of **c** =0.05.
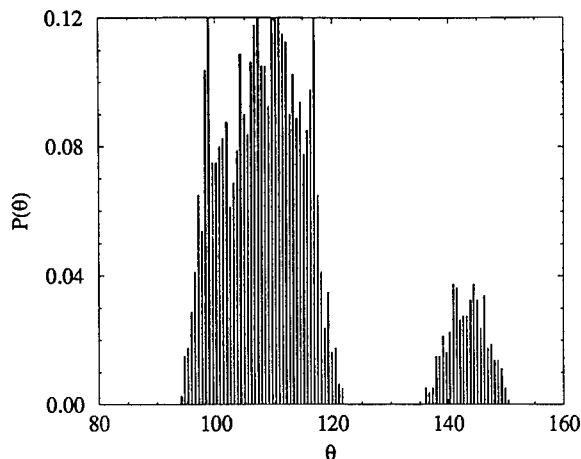
Figure 9: Bond distribution function P(θ) of the three-fold coordinated sites in the case of **c** =0.01. Compare this distribution with that reported in fig.4 (dotted line), relative to the case c=0.18.

ture udergoes to a crystalline-to-amorphous transition, as testified by the pair correlation function reported in fig.5.

CONCLUSIONS

The present results can be resumed in the following points:

1. Semi empirical TBMD, in the used parametrization [15] allows, a qualitative reproduction of the a-C structures. We have underlined the inadequacy of the parametrization to correctly estimating the concentration of four-fold coordinated sites in the high density ($\delta = 3$ gr cm^{-3}) systems, when they originated by fully disordered structures (e.q. liquid). In the case of structures which undergo a progressive disordering process, in turn, the used TB parametrization is able to retain into the system a large fraction of sp^3 sites ;

2. The EDOS resulting from the calculations qualitatively agrees with experimental data [10], although the relative peak intensities are affected by the problem pointed out above;

3. The optical (Tauc) gap decreases linearly with the increase of the sp^2 fraction, in agreement with [27];

4. The crystal-to-amorphous transition, triggered by the introduction of dissociated Frenkel pairs in a diamond lattice, occurs continuosly. This transition is induced by the presence of large local strains associated to three-fold coordinated sites which retain the sp^3 hybridization (dangling bonds). These structures, characterized by a "quasi" planar configuration, transform into "true" sp^2 hybrids. It can be conjectured that this process occurs as soon as the local concentration of strain exceeds some instability threshold;

5. A complete a-C system has been found when the concentration of introduced Frenkel pairs exceeds the value of 0.15.
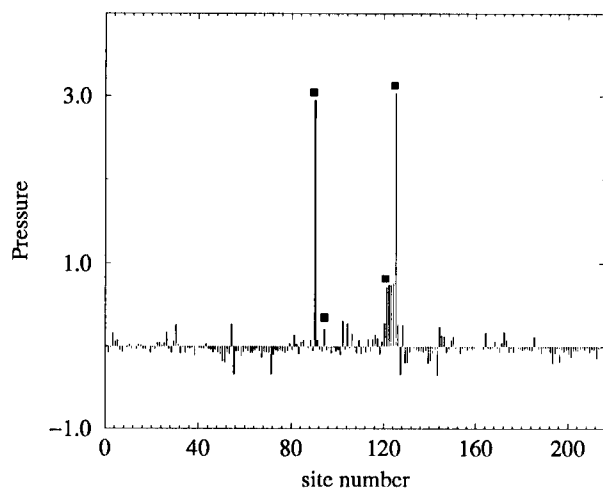
Figure 10: Local pressure on the atomic sites in system with $c = 0.005$. Pressures are reported in eV $\mathring{A}^{-3}$. Full squares indicate the three-fold coordinated sites in the structure.

REFERENCES

1. G.Galli, R.M.Martin, R.Car and M.Parrinello, Phys. Rev. Lett. **62**, 555 (1989).

2. C.Z.Wang, K.M.Ho, C.T.Chan, Phys.Rev.Lett. **70**, 611 (1993).

3. D.A.Drabold, P.A.Fedders, P.Stumm, Phys.Rev.B **49**, 16415 (1994).

4. C.Z.Wang, K.M.Ho, Phys.Rev.B. **50**, 12429 (1994).

5. Th.Frauenheim, P.Blaudeck, U.Stephan and G.Jungnickel, Phys. Rev. B **48** 4823 (1993).

6. N.A.Marks, D.R.McKenzie, B.A.Pailthorpe, M.Bernasconi and M.Parrinello. Phys. Rev. Lett. **76**, 768 (1996).

7. N.A.Marks, D.R.McKenzie, B.A.Pailthorpe, M.Bernasconi and M.Parrinello, Phys. Rev. B **54**, 9703 (1996).

8. V.Rosato, M.Celino and L.Colombo, Comp. Mat. Sci. in press

9. J.Tersoff, Phys.Rev. B **37**, 6991 (1988).

10. J.C.Lascovich, A.Santoni, V.Rosato, submitted to Phys.Rev.B.

11. F.Li, J.S.Lannin, Phys.Rev.Lett. **65**, 1905 (1990).

12. P.H.Gaskell, A.Saeed, P.Chieux, D.R.McKenzie, Phys.Rev.Lett. **67**, 1286 (1991).

13. K.W.Gilkes, P.H.Gaskell and J.Robertson, Phys. Rev. B **51**, 12303 (1995).

14. C.H.Xu, C.Z.Wang, C.T.Chan and K.M.Ho, J. Phys. Cond. Matter **4**, 6047 (1992).

15. L.Colombo in: Annual Review of Computational Physics IV, edited by D.Stauffer (World Scientific, Singapore, 1996) p.147.

16. I.Kwon, R.Biswas, C.Z.Wang, K.M.Ho, C.M.Soukoulis, Phys.Rev.B **49**, 7242 (1994).

17. Y.Limoge, A.Barbu, Phys. Rev. B **30**, 2212 (1984).

18. L.Colombo, D.Maric, Europhys.Lett. **29**, 623 (1995).

19. C.Massobrio, V.Pontikis and G.Martin, Phys. Rev. Lett. **62**, 1142 (1989).

20. C.Massobrio and V.Rosato, Solid State Phenomena **23&24**, 147 (1992).

21. R.L.McGreevy and L.Putsztai, Mol. Sim. **1**, 359 (1988).

22. J.K.Walters and R.J.Newport, Phys. Rev. B **53**, 2405 (1996).

23. J.K.Walters, K.W.R.Gilkes, J.D.Wicks and R.J.Newport, J. Phys.: Condens. Matter **9**, L457 (1997)

24. V.Rosato, J. C. Lascovich, A. Santoni, L. Colombo, accepted for publication on J.Non Cryst. Sol.

25. J.Tauc in Amorphous and Liquid Semiconductors, edited by J.Tauc (Plenum, London 1974) p.176.

26. J.Robertson, Phil. Mag. B **76**, 335 (1997).

27. J.Robertson, Phil. Mag. B **66**, 199 (1992).

28. J.Bernholc, A.Antonelli, T.M.Del Sole, Y.Bar-Yam, S.T.Pantelides, Phys. Rev.Lett. **61**, 2689 (1988).

29. T.Egami, K.Maeda, V.Vitek, Phil.Mag.A **41**, 883 (1980).

COVALENT LIQUIDS: TIGHT BINDING SIMULATION VERSUS EXPERIMENTAL RESULTS

J.-P. GASPARD *, C. BICHARA **, J.Y. RATY *
° Université de Liège, Institut de Physique, B5, 4000-Sart-Tilman, Belgium, jp. gaspard @ulg.ac.be
** Centre de Thermodynamique et de Microcalorimétrie - CNRS, 26, rue du 141ème R.I.A., 13003 Marseille, France

ABSTRACT

We show that a simple tight binding model with a repulsive potential describes the Peierls distortions in covalent systems and the well-known octet rule. The existence and the intensity of the Peierls distortion is mainly related to the hardness of the repulsive potential as demonstrated both by theoretical calculations and by the experimental systematic analysis of liquid structures. In particular, As is threefold coordinated and Sb is sixfold coordinated in the liquid, the qualitative difference is explained by the ratio of the distortion energy ΔE to the thermal energy $k_B T$. The $As_x Sb_{1-x}$ alloys show continuously varying average coordination numbers showing: the semiconductor-metal transition is continuous with concentration.

In addition, we illustrate in the case of liquid Se that, tight binding Monte Carlo simulations are able to describe quantitatively the structure of liquid elements provided the Van der Waals potential is added.

INTRODUCTION

The first aim of this paper is to show that a simple tight-binding model is able to explain the crystallographic structures of phases that are driven by electronic instabilities as well as their phase transitions. In the first part, we address the general question of the stability of covalent systems, whose structure is essentially governed by directional p bonding, resulting in open structures.

The Peierls distortion is an electronic instability that produces an alternation of short and long bonds following a rule that depends on the band filling ratio. The process is largely independent of the details of the model. In the simplest case, the tight binding method allows to obtain simple - sometimes analytic- expressions or simple relations between the quantities of physical interest e. g. between the total energy and the compressibility or the Einstein oscillator frequency. We stress upon the fact that the repulsive potential plays an important role, and we show that its variation with distance is a relevant parameter that determines the type of stable structure.

Elements of the columns V, VI and VII of the Mendeleiev table satisfy the $Z = 8 - N$ rule, with a limited number of exceptions, i. e. the coordination number Z is equal to eight minus the number of s and p valence electrons. In particular, As and Sb atoms are three-coordinated in corrugated planes while Se and Te atoms are two-coordinated in helicoidal chains. These structures are described as Peierls distortions of the simple cubic structure as shown in Fig. 1.

We have shown that these structures result from an electronic instability (or symmetry breaking mechanism) that shortens and elongates the bond length with a periodicity that depends on the band filling ratio [1]. Simultaneously, an energy gap appears at the Fermi energy and the shift of the occupied states close to the Fermi energy towards more bonding energies produces an energy gain. This energy gain is counterbalanced by an increase of the repulsive energy, and it is understandable that the balance depends on the hardness of the repulsive potential. An interesting question is the robustness of this distortion with respect to pressure and temperature. It is well known that the pressure reduces the long bond distance and so the Peierls distortion is reduced up to disappearance. It is clearly understandable that the repulsive term plays an important role in the existence or the absence of a Peierls distortion. In particular we have shown that smooth repulsive potentials leads to distorted structures whilst the "hard sphere" case does not. In this paper, we discuss the stability of the distortion with respect to the temperature, in framework of a simple model. From the experimental side, in $As_x Sb_{1-x}$ alloys an extensive set of diffraction data exists that are in full agreement with the model predictions [2].
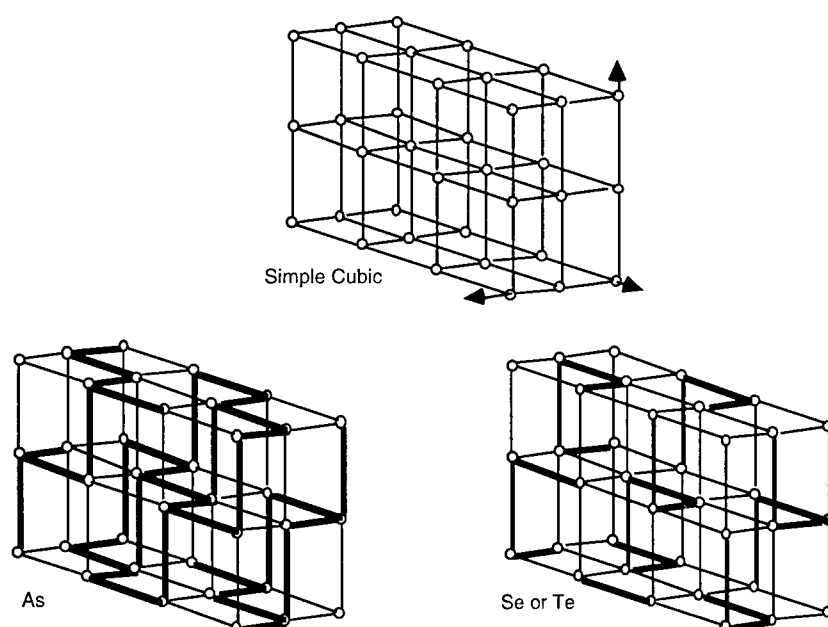
Figure 1. Simple cubic structure and structures of As, Se and Te. Short bonds are indicated by heavy lines and long bonds by light lines.

In a second part, we show that a more elaborate tight binding model provide a quantitative description of the structure of matter under extreme conditions of temperature (and pressure). The example considered here is liquid selenium whose structure and electronic properties from the melting point to supercritical conditions have been extensively studied in the past few years [3]. One of the main reason for this interest is that the high temperature liquid phase is metallic [4], with an average coordination number *lower* than that of the liquid just above the melting point, which is semi conducting, like crystalline and amorphous selenium. This is a rather unexpected behaviour, as the semi conductor to metal transition that occurs in the Peierls distorted covalent elements when applying temperature or pressure is usually accompanied by an *increase* of the number of first neighbours. A well known example for this behaviour is tellurium [5]. Moreover, recent EXAFS measurements [13] have shown that this transition is accompanied by a significant broadening of the first neighbour shell towards small distances, which is consistent with an increase of the number of one-fold coordinated atoms (ends of chains or molecules) in the melt. Various theoretical explanations have been proposed, using either simple tight binding models [6] to analyse the electronic properties or ab initio techniques [7] to study the crystalline and liquid phases. The computer simulation approaches, the only able to deal quantitatively with the physics of liquids, face a challenging problem due to the nature of the bonding in selenium. The bonds within the chains are clearly covalent with an energy of the order of 2eV/atom, whereas the bonds between the chains are an order of magnitude weaker, an energy range where dispersion forces are no longer negligible and must be taken into account, either directly by adding a Van der Waals term in the semi empirical models, or by including corrections beyond the standard LDA in ab initio approaches. In the liquid state, specially at high temperature, the density is much lower than in the crystal (typically .026 atom/Å^3 vs .036 atom/Å^3) and this problem becomes crucial. Although it has been possible to simulate liquid selenium at the melting point with a very crude model neglecting the dispersion terms [8], such an attempt leads to unphysical large concentration fluctuations in the

high temperature conditions, where the density is low. The present work includes of a Van der Waals term, in addition to the tight-binding hamiltonian, of course at the expense of an increase of the number of parameters.

QUALITATIVE TIGHT-BINDING

In this section ,we present a simple statistical model that accounts for the variation of the coordination number with temperature. The model is as follows: we assume that the atoms are sitting on a simple cubic lattice but the distance is artificially modified by varying the interatomic resonance integral β. The integral β can even be put equal to zero. This a simple and useful technique for changing the distance without changing the angles. The energy is calculated and equilibrated along any pair of neighbouring atoms. The attractive energy is treated in a second moment approximation. In a simple cubic structure and including only $pp\sigma$ interactions, the three directions of space are decoupled and the total energy writes, per direction:

$$E_{tot} = -\sqrt{Z}\,\beta\,\frac{1}{r^q} + \frac{Z}{2}\,V_0\,\frac{1}{r^p} \qquad (1)$$

Z is the coordination number equal to 0, 1 or 2 in a given direction. The first term is the attractive band term and the second term describes a pairwise addition repulsive contribution. Let is recall that $p > q$ to insure structural stability [1]. For $p > 2q$ the structure with $Z = 2$ (per direction) is the most stable while for $p < 2q$, an electronic instability occurs and $Z = 1$. Using formula (1), it is easy to show that the total energy, at equilibrium, is proportional to $Z^{2\frac{p-2q}{(p-q)}}$. Formula (1) or its extension to two different distances r_s and r_l has an important artefact: at zero pressure, the equilibrium distance of the long bonds is infinitely long i. e. the bond is broken. In order to get a finite equilibrium distance r_l, one has either to fix the total volume or to take the dispersion forces into account as we do in the next section.

Roughly, the larger the number of core electrons, the harder the repulsive potential and the larger the p value. This is illustrated by the case of As_xSb_{1-x} alloys. This series is interesting as it varies continuously the hardness parameter within the interval $[p_{As}, p_{Sb}]$.

On the other hand, the leading entropy term, treated as a collection of Einstein oscillators, is proportional to $k_B \ln (V/V_0)$ where V_0 is some reference volume and V is the volume of the vibration ellipsoid. The sampled volume increases at high temperature in the case $p < 2q$ due to the anharmonicity of the effective potential. As a consequence, temperature increases the average coordination number (Fig. 2 and, 3 and 4).

As a result of a series of neutron diffraction experiments, we find out that there is an increased number of coordination with the concentration in the heaviest element (Sb here), from 1 to 2 per direction, or from 3 to 6 for the three directions of space (Fig. 3). This is clearly illustrated by the case of $As_x Sb_{1-x}$ alloys. The crystalline structures of arsenic and antimony are isostructural. the two structure differs from the intensity of the Peierls distortion which decreases from the top to the bottom of the fifth column of the periodic table. This result in variation of the rhombohedral angle and a relatively larger separation between corrugated planes of three-coordinated atoms in As, compared to Sb. The strength of the Peierls distortion decreases with the atomic number or, in other words, the energy gain of a Peierls distortion is smaller when going down the periodic table. At high temperature the thermal energy can be equal to the Peierls distortion energy and the distortion is washed out. Neutron diffraction experiments have shown that the coordination number increases gradually from 3 in liquid As up to 6.3 in liquid Sb [2]. Figure 3 shows the variation of the coordination number with the concentration, just above the melting temperature. The evolution can be understood as a destruction of the Peierls distortion. Monte Carlo simulations of the liquid phase of As and Sb using a classical tight-binding description of bonding with a pairwise repulsive potential $V = C\,r^p$ have shown the importance of the p exponent of the interatomic repulsion [3, 6]: As has a larger value of p that Sb. The structure of liquid $As_x Sb_{1-x}$ alloys can then be easily simulated using the geometrically averaged values of the repulsive

potentials of the individual elements. By using the entropy approximation and the lattice model described here above, we can estimate the behaviour of the alloys with temperature.
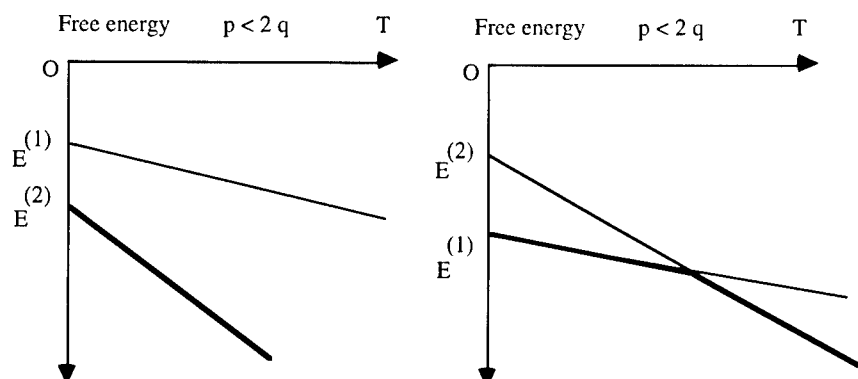
Figure 2. Free energy (schematic), as a function of temperature, of 1-coordinated and 2-coordinated structures for hard spheres (left) and soft spheres (right).
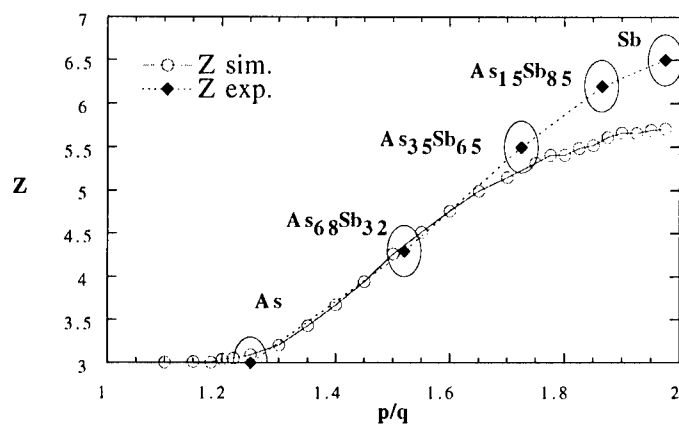
Figure 3. Coordination number at 980 K, as a function of the hardness parameter p/q of the repulsive potential.

Fig. 4 shows the comparison between the coordination number, measured as the total number of short bonds around an atom, for the lattice tight binding calculation and for the neutron diffraction data, as a function of the p/q ratio at the temperature of 980K. Of course, the result of our model cannot fully agree with the experimental curve since the maximum coordination number is, by construction, limited to six in the model. A more detailed analysis of the coordination number is given on Fig. 4. Starting from 100% of three coordinated atoms in liquid As, we see that the four coordinated atoms become majority for p/q values between 1.4 and 1.55, and then the same for five coordinated atoms up to p/q equal to 1.7. Coordination six is then predominant for higher p/q values.
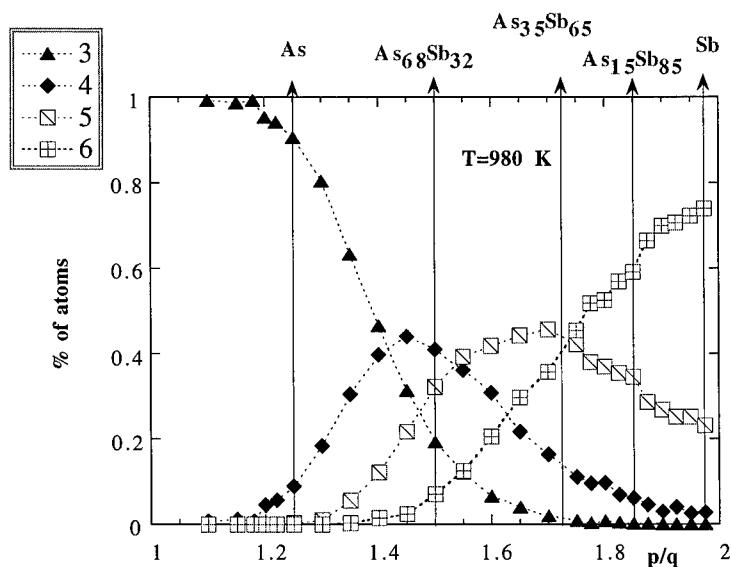
Figure 4. Probability distribution of 3, 4, 5 and 6-coordinated atoms as a function of the hardness parameter p/q of the repulsive potential.

This example shows clearly how our simplified tight-binding model, combined with an estimation of the entropy, describes qualitatively the evolution of the most simple Peierls distorted systems as a function of the hardness of the repulsive potential or the temperature.

QUANTITATIVE TIGHT-BINDING AND MONTE-CARLO SIMULATION

In a second step, we put forward a quantitative method for the accurate determination of the structure factor of liquids with two different bonding mechanisms (covalent and Van de Waals). In the tight binding method, which is semiquantitative, we anchor the parameters to different measurable quantities such as the total energy, the equation of state V(P) and the band structure obtained by a LMTO technique. Both energetic and spectroscopic parameters are used in the fitting procedure.

The model

The total energy of the system is the sum of three terms : a band structure term, an empirical pairwise additive repulsive term and a third term, which is also pairwise additive, includes the dispersion forces. The total energy writes :

$$E_{tot} = E_{band} + E_{rep} + E_{VdW} \tag{2}$$

$$E_{band} = \int_{-\infty}^{E_F} E \ n(E) \ dE \tag{3}$$

457

$$E_{rep} = \frac{1}{2} V_0 \sum_{i,j} \frac{1}{r^p} F_1(r) \tag{4}$$

$$E_{VdW} = \frac{1}{2} C_6 \sum_{i,j} \frac{1}{r^6} (1 - F_2(r)) \tag{5}$$

The band energy is calculated using a fourth moment approximation of the density of states. Although crude, this approximation retains the essential physics of the system (the quantum nature of the chemical bond, the directional character of the covalent bond, the gap at the Fermi level for a Peierls distorted element). More details are given in [8]. The method allows relatively fast computations on large systems (up to some thousands of atoms on a workstation). Moreover, a more accurate description of the DOS is irrelevant considering the crude approximations made on the other terms. The damping terms $F_i(r) = [1 + \exp(\frac{r - \alpha_i}{\delta_i})]^{-1}$ ensure a smooth cancellation of the first two terms at the cut-off distance. The damping of the Van der Waals term $(1-F_2(r))$ allows a smooth transition between the long range regime, where the approximate expression of the dispersion term (5) is valid, and the short range regime where the covalent bonding term given by (3) and (4) is physically dominant. The parameters (α_i, δ_i) of these damping functions as well as the parameters defining the hopping integrals $\beta_\lambda = \beta_\lambda{}^\circ \exp(-q_\lambda r)$ (β_λ stands for the sso, spo, ppo and ppπ interactions considered here) and the repulsive term must be carefully fitted in order to obtain realistic results.
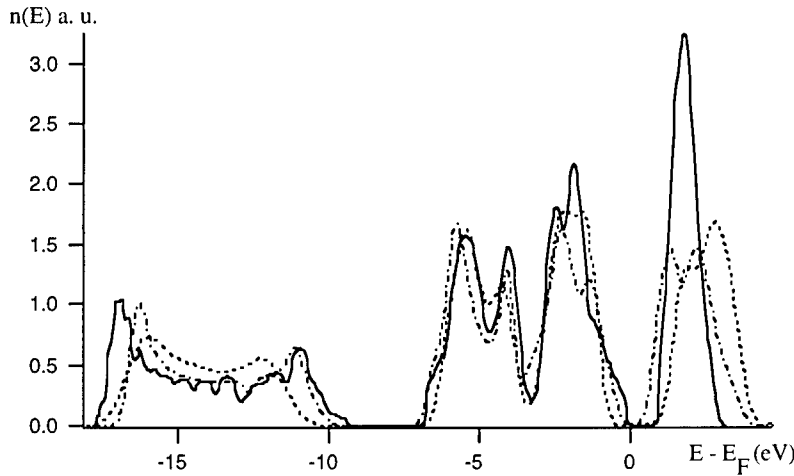
Figure 5 : Electronic density of states of SeI at 0 GPa. Full curve : tight binding; dash-dotted curve : FP-LMTO; dotted curve : ref [12].

The fitting technique proceeds in two steps. First, the tight binding parameters p_0, q_λ are obtained by fitting the first moments of the electronic densities of states (DOS) that have been calculated using the FP-LMTO method [9] on different crystalline trigonal (Se^I) selenium structures in the pressure range 1-10 GPa. The interest of considering various pressures is that it provides one with the distance dependence of the hopping integrals, and hence q_λ . The structural parameters used for these calculations are the experimental ones [10, 11]. The resulting DOS for

Selenium at 0 GPa is plotted on Fig. 5, compared with our FP-LMTO calculations and the calculations of Schlüter et al. [12]. The tight binding density of states is calculated using 40 exact moments. In a second step, the remaining parameters (repulsive term, damping functions) are calculated by fitting the experimental equations of state [10]. The structural parameters for the phases Se^{II} and Se^{III} are not available, so we restrict ourselves to the Se^{I} and Se^{IV} phases. The experimental curve for the Se^{IV} phase has been calculated assuming a direct transition pressure between Se^{I} and Se^{IV} at 25 GPa. The resulting energy versus reduced volume curves are plotted in Fig. 6. It can be seen that the agreement is rather good for the Se^{I} phase around its equilibrium volume at 0 GPa, and less satisfactory for high pressures. As our simulations are done in the liquid state at low densities, these problems are expected to remain negligible.
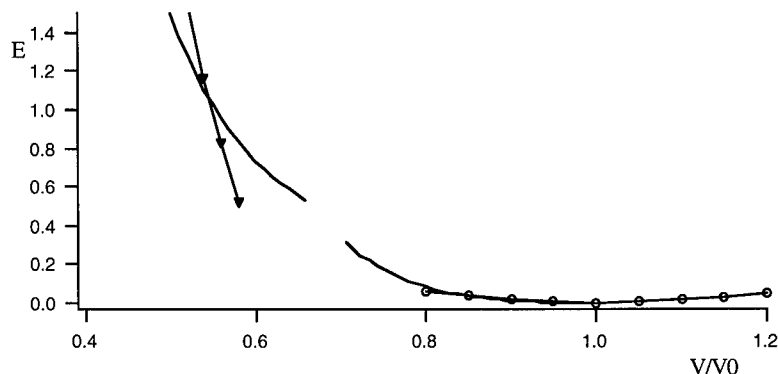
Figure 6. Total energy vs reduced volume. (———) experimental [10], circles and triangles : tight binding fit.

With the parameters determined in this way, (N,V, T) Monte Carlo simulations have been performed under the density and temperature conditions indicated in Table 1. Apart from the lowest temperature, they correspond to the conditions (P = 600 bars) under which the EXAFS experiments of Soldo et al. [13] have been performed. The results presented here have been obtained for a system size of 144 atoms, and up to 10000 Monte Carlo steps / atom have been necessary to reach equilibrium.

Results

The pair correlation functions g(r) and structure factors S(q) calculated under these conditions are plotted on Fig. 7 and 8 respectively. The structure factor at the lowest temperature is in very good agreement with the experimental results of [14]. Moreover, the temperature dependence of the structure factor displays the characteristic features observed in [1] : a gradual decrease of the first peak around 2 Å^{-1}, accompanied by an overall damping of the large q oscillations with increasing temperatures.

The detailed analysis of the distribution of the coordination shells (Fig. 9 and Table 1) shows that the average coordination number $<N_1>$, that decreases with increasing temperature, is an average of various fractions of 1-fold, 2-fold and 3-fold coordinated atoms. Table 2 presents upper and lower bounds on the measured quantities : they have been calculated for two different cut-off values (2.60 Å and 2.85Å) that are reasonable bounds for this quantity. The amount of 2-fold coordinated atoms is approximately constant (around 60%), whereas the number of 1-fold coordinated atoms gradually increases and the number of 3-fold coordinated atoms decreases with increasing temperature. The bond of one-fold coordinated atoms (ends of chains) is shorter than the bonds between 2-fold and 3-fold coordinated atoms, hence the gradual shift of the first peak position. This shift, clearly apparent on Fig. 7, can be measured by calculating the average of the first peak position ($<R_1>$: normalised first moment of the first peak).

459
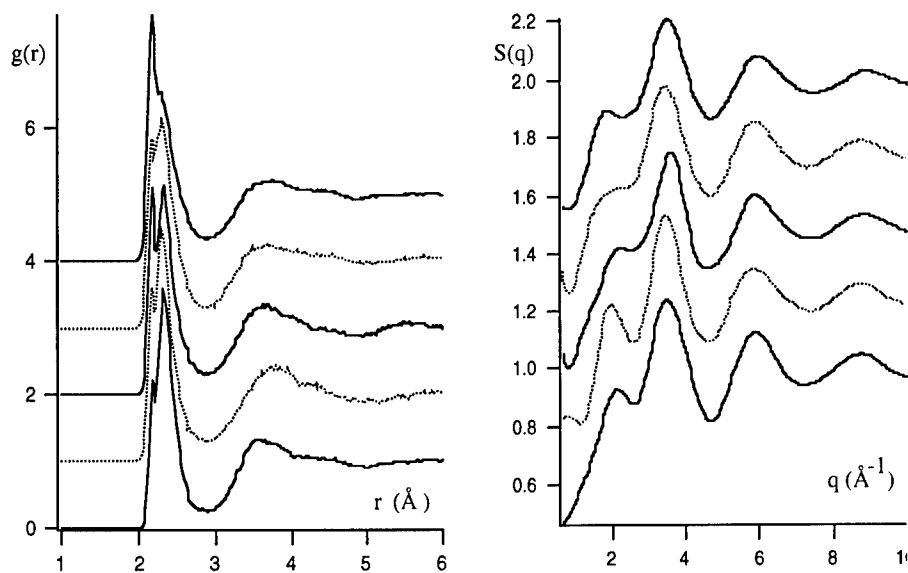
Figures 7 and 8. Pair correlation function (g(r), left) and structure factor (S(q), right) of liquid selenium at 890, 1075, 1180, 1350, 1510°C (from bottom to top).
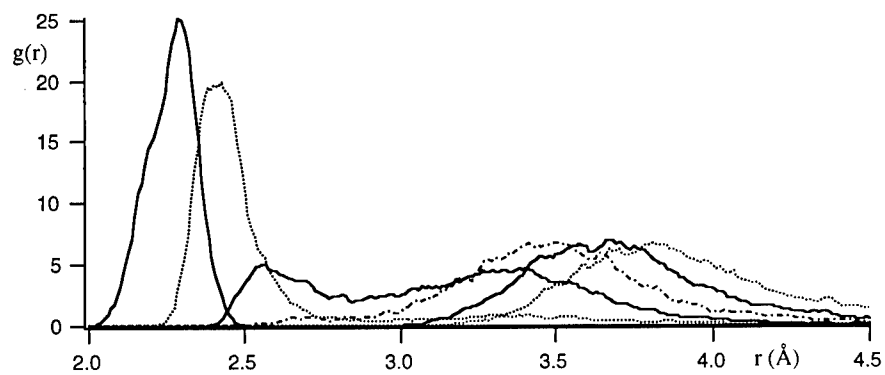
Figure 9. Partial pair correlation function r^2 g(r) corresponding to the six first neighbours (left to right).

Table 1 - Upper and lower bounds on the average coordination number <N1>, the distribution of the coordination and the average first peak position <R₁>, calculated with cut-off distance at 2.60Å and 2.81Å.

Temperature (°C)	Atomic density (Å^{-3})	$<N_1>$	$\%CN_1$	$\%CN_2$	$\%CN_3$	$<R_1>$ (Å)
890	.0259	1.90-2.11	23-15	64-61	13-24	2.35-2.38
1075	.0249	1.85-2.07	34-18	54-59	12-23	2.34-2.38
1180	.0244	1.77-2.00	33-24	57-54	11-22	2.34-2.37
1350	.0236	1.82-2.06	29-19	60-58	11-23	2.34-2.38
1510	.0229	1.67-1.90	40-28	53-54	7-18	2.32-2.36

The tight binding calculations predict a metallic behaviour on the basis of the electronic densities of states calculated at 1180°C and 1510°C. A more elaborate electronic structure calculation on characteristic configurations would certainly be more reliable but is beyond the scope of this preliminary work. Nevertheless, it turns out that the electronic density of states near the Fermi level depends crucially on the (broad) distribution of third neighbours (distances between 2.5 and 3.5 Å) that appears in the liquid state, as can be seen on Figure 10. This part of the real space atomic distribution that corresponds to the gap between the first and second peaks of the pair correlation function, is rather difficult to sample correctly, even with the rather long simulation times used here. In this region, two competing effects take place. The effect of the temperature that tends to fill the gap between the first and second peaks of g(r) is counter balanced by the effect of the decreasing density that tends to set the chains (or segments of chains) apart, as confirmed by the tendency to lower the number of three-fold coordinated atoms with increasing temperature. It would probably be more reliable to let the system adapt its density to the temperature by using a more time consuming N, P, T algorithm for the Monte Carlo simulations.

n(E) a. u.
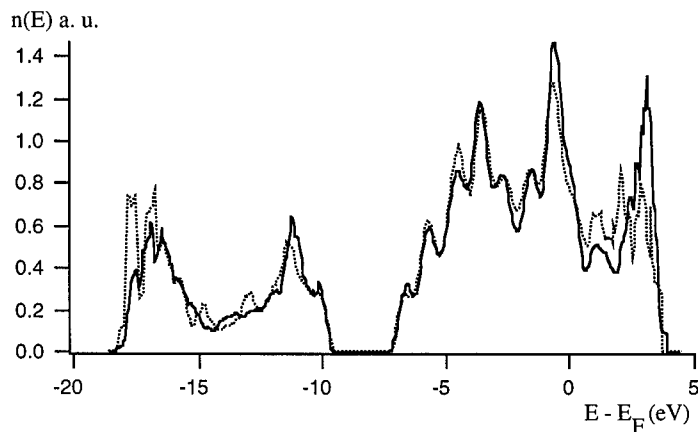
Figure 10. Electronic density of states of liquid Se at 1075° (full curve) and 1510°C (dash-dotted curve).

Liquid selenium is intermediate between a purely covalent and a molecular system, and the

efficient sampling of the phase space of a system where two different energy scales coexist is not a trivial matter and requires further work.

CONCLUSIONS

Although semi-quantitative, the tight binding method is important not only for the qualitative understanding of bonding mechanisms but also to obtained a detailed and accurate description of the liquid structures, from which various parameters can be calculated. In the case where an electronic instability occurs, the dispersion (Van der Waals) energy has to be taken into account in a parametrized form in order to get quantitative results for the liquid structure. The advantage of this method over the ab initio technique is a reduced amount of computing time that allows either to treat thousand of atoms or to have good statistics on smaller samples.

REFERENCES

[1] Gaspard J. P., Pellegatti A., Marinelli F. and Bichara C., Phil. Mag. B, to be published
[2] Raty, J. Y., Gaspard J. P., Bellissent R. and Ceolin R., NCM7 Proceedings, to be published
[3] Inui M., Noda T., Tamura K. and Li C., J. Phys. Condens. Matter, **8**, 1996, p. 9347-51 .
[4] Hosokawa S. and Tamura K., J. Non-Cryst. Solids, **117-118**, 1990, p. 482-92
[5] Bichara C., Raty J.-Y. and Gaspard J.-P., Phys. Rev. B53, **1**, 1996, p. 206-11.
[6] Koslowski T., J. Phys. Condens. Matter, **9**, p1997, . fi .
[7] Kirchhoff F., Gillan M. J., Holender J. M., Kresse G. and Hafner J., J. Phys. Condens. Matter, **8**, 1996, p. 9353-7 and, Kresse G., Furthmuller J. and Hafner J., Phys Rev. B50, 1994, p. 13181.
[8] Bichara C., Pellegatti A. and Gaspard J.-P., Phys. Rev. B49, **10**, 1994, p. 6581-6.
[9] M. Methfessel, Phys. Rev. B 38, 1988, p. 1537.
[10] Krüger T. and Holzapfel W. B., Phys. Rev. Lett. 69, **2**, 1992, p. 305-7.
[11] Akahama Y. Kobayashi M., and Kawamura H., Phys. Rev. B47, **1**, 1993, p. 20-6.
[12] Schlüter M., Joannopoulos J.D. and Cohen M.L., Phys. Rev. Lett. 33, **2**, 1974, p. 89-91.
[13] Soldo Y., Hazemann J.-L., Aberdam D., Inui M., Tamura K., Raoux D., Pernot E., Jal J.-F., Dupuy-Philon J., Phys. Rev. B., to be published
[14] Bellissent R. and Tourand G., J. Non-Cryst. Solids, **35-36**, 1980, p. 1221

STRUCTURAL AND ELECTRONIC PROPERTIES OF *a*-GaAs: A TIGHT-BINDING–MOLECULAR-DYNAMICS–ART SIMULATION

LAURENT J. LEWIS[a]
Département de physique et GCM, Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, Québec, Canada H3C 3J7
NORMAND MOUSSEAU[b]
Computational Physics — Faculty of Applied Physics, Technische Universiteit Delft, Lorentzweg 1, 2628 CJ Delft, The Netherlands

ABSTRACT

By combining tight-binding (TB) molecular dynamics (MD) with the recently-proposed activation-relaxation technique (ART), we have constructed structural models of *a*-GaAs and *a*-Si of an unprecedented level of quality: the models are almost perfectly four-fold coordinated and, in the case of *a*-GaAs, exhibit a remarkably low density of homopolar bonds. In particular, the models are superior to structures obtained using melt-and-quench TB-MD or quantum MD. We find that *a*-Si is best described by a Polk-type model, while *a*-GaAs resembles closely the mechanical model proposed by Connell and Temkin, which is free of wrong bonds. In this paper, the structural, electronic, and dynamical properties of *a*-GaAs based on this approach will be reviewed, and compared to experiment and other structural models. Our study provides much-needed information on the intermediate-range topology of amorphous tetrahedral semiconductors; in particular, we will see that the differences between the Polk and Connell-Temkin models, while real, are difficult to extract from experiment, thus emphasising the need for realistic computer models.

INTRODUCTION

The energetics, and thus the physical properties, of multicomponent materials is notably difficult to describe accurately using simple empirical potentials because of the large number of interactions involved. Because of this, with perhaps the exception of water and SiO_2 (and related chalcogenide glasses), very little effort has been spent in developing such potentials. In light of this, model calculations based on a tight-binding (TB) or *ab-initio* description of the interactions are no luxury but, rather, a much-needed first step in the study of innumerable, and important, multicomponent systems.

However, the computational cost of *ab-initio* calculations, and to a lesser extent TB, increases rapidly with system size. This becomes a serious problem for systems which exhibit little (or no) symmetry, such as nano-structures, surfaces or disordered (chemically or topologically) materials, where large models are required and for which, therefore, the number of relevant configurations increases rapidly. One possible solution to this problem is the use of different approaches at various stages of a calculation. Thus, for instance, a first-approximation structural model for a particular material might be generated using a set of classical empirical potentials, which can then be improved or confirmed *a posteriori* through further relaxation using more appropriate (semi-empirical or first-principles) interactions. This is the philosophy we adopt here.

Mixed approaches have been used with some success in the study of disordered materials, such as *a*-Si, *a*-C, etc.[1]. Although leaving open *a priori* the possibility that the empirical potentials bias the results in a significant manner, these approaches probably provide, at present, and until computers and methodology improve substantially in speed, the most efficient method for optimising complex structural models. *A posteriori*, the structural models can be checked against experimental data and/or other models, when available.

We give here an example of a mixed-approach calculation as applied to the study of amorphous silicon (a-Si) and amorphous gallium arsenide (a-GaAs): A starting point structure is constructed using empirical potentials and the activation-relaxation technique, then relaxed using TB forces in a molecular-dynamics (MD) simulation. The resulting structural models are found to be superior to structures obtained using the conventional melt-and-quench, TB- or quantum-MD approaches.

The standard Polk continuous-random-network (CRN) model [2] is often considered to represent the ideal structural model for amorphous elemental semiconductors such as a-Si, a-Ge and fourfold a-C. For compound semiconductors, such as the III-V GaAs, the Polk model is likely not ideal: it contains a macroscopic proportion of odd-membered rings which are absent in the crystal. The presence of such rings in a binary mixture such as GaAs would necessarily result in the presence of both hetero and homopolar (wrong) bonds. At best, the Polk model would still possess at least 12% of wrong bonds [3]. These are very expensive in GaAs because of the partly ionic character of the interactions. This cost, however, must be balanced against that for elastically deforming the network to eliminate the bonds between like atoms.

Connell and Temkin (CT) [4] have demonstrated that it is quite possible to construct a CRN *without* odd-membered rings, thus allowing for the construction of chemically-ordered binary semiconductor models. The (mechanical) CT model, however, suffers from a large surface and cannot address the question of the elastic cost arising from the elimination of wrong bonds, or, equivalently, odd-membered rings. With computer models, it is possible to suppress surfaces, through the use of periodic boundary conditions, but eliminating wrong bonds in a material such as a-GaAs, which is relatively weakly ionic (compared, e.g., to SiO_2) [5], is difficult because the timescale for chemical ordering to take place is longer than can be afforded in MD simulations with TB or first-principles interactions.

Experimentally, the situation is not much clearer. Because the interatomic distances for the different pairs (Ga-Ga, As-As and As-Ga) are almost identical, probes such as EXAFS yield very little information about the local environment of each type of atoms. Indirect measurements conclude to the near absence of homopolar bonds in this material but, as we will show here, these conclusions are not warranted by our simulations, which indicate that wrong bonds affect very little average quantities.

The solution we adopt in this work proceeds from the mixed approach discussed earlier: First, we use the activation-relaxation technique (ART) of Barkema and Mousseau [6, 7], coupled with empirical (but modified) Stillinger-Weber potentials, to prepare two 216-atom models with periodic-boundary conditions, corresponding to the Polk (P) and Connell-Temkin (CT) CRN's, respectively. Second, *both* models are relaxed with the Goodwin-Skinner-Pettifor TB potential for Si [8] and the Colombo-Molteni TB for GaAs [9]. We find that a-Si is best described by a Polk-type model, while a-GaAs resembles closely the CT model. The differences between the two models, while real, are however difficult to extract from experiment, thus emphasising the need for accurate computer models.

EMPIRICAL APPROACH

Satisfactory models of a-GaAs are extremely difficult to construct for two reasons: (i) There exists, to our knowledge, no empirical potential which can provide a realistic description of the atomic interactions in this material; one therefore has to resort to TB or *ab initio* forces. (ii) As mentioned earlier, a-GaAs is (relatively) weakly ionic and the timescale associated with chemical ordering is therefore beyond the reach of conventional MD simulations (in contrast to, e.g, SiO_2), especially if TB or *ab initio* forces are to be used.

As a first step, prior to relaxing using TB forces, and in order to generate computer structures which are akin to the Polk and Connell-Temkin CRN's, we carried out ART calculations (see below) using modified Stillinger-Weber potentials, where the three-body

term is enhanced by a factor of 2 (i.e., $\lambda = 42$ compared with 21 in Ref. [10]; the standard SW parameters cannot lead to a 4-fold CRN [6]). Since all 216 atoms are equivalent in this model, the resulting structure corresponds to the Polk model and *does* possess odd-membered rings; the model will be referred to as CRN-P. To eliminate (or at least minimise) the number of odd-membered rings, and thus obtain a CT-type CRN (hereafter referred to as CRN-CT), the atoms are labeled either A or B, and the model relaxed subject to an additional short-range repulsion between like-atoms (keeping full symmetry between A and B):

$$E_{rep} = \sum_{<ij>} A_{ij}\epsilon \left[1 + \cos\left(\pi \frac{r_{ij}}{s_{ij}}\right)\right] \tag{1}$$

where the sum is over all pairs of atoms, ϵ is the Stillinger-Weber energy scaling factor, $s_{ij} = 3.6$ Å, and $A_{ij} = 1.2$ for like particles and zero otherwise.

The optimisation of the two CRN's was performed using ART, a Monte Carlo procedure which forces the system to "jump" from one local minimum to another via a saddle point. ART reduces the configurational energy landscape to a discrete (but infinite) network of local minima connected by paths going through first-order saddle-points. Since the algorithm is local *in the configurational energy landscape*, ART proceeds through a random-walk on this network. New moves are accepted *modulo* a Boltzmann factor whose fictitious temperature follows a simulated-annealing procedure. Although local in the configurational energy landscape, ART is non-local in real space, leaving full freedom for the type of excitation and atomic moves involved in the passage from one minimum to another. Thus, collective relaxation mechanisms can take place without restriction, enhancing the efficiency of the algorithm. Full details of ART can be found in Refs. [6, 7].

In both cases, the initial configuration consisted of 216 atoms placed at random in a cubic box of appropriate density. ART was then invoked repeatedly until the configurational energy converged to a state from which it is difficult to escape for a fictitious temperature of 0.25 eV. In the case of CRN-P, the relaxation is purely topological, and a well-relaxed configuration can be obtained within a day or so on a SGI-R8000 workstation. For CRN-CT, chemical ordering requires longer-range atomic displacements and, in practice, runs took on the order of a week to converge.

There are concerns, with such optimisation schemes, that final structures might be partly crystalline (rather than amorphous). The method was tested on a 64-atom unit cell and in all cases, crystallisation was found to occur. However, crystallisation was never observed on 216-atom unit cells, which converged, based on visual inspection, to "smooth" amorphous structures. The reason for this difference in behaviour, which actually demonstrates the power and efficiency of the method, is that ART is local in configuration-energy space; for a small system, the probability of finding the global minimum (the crystalline state) is relatively large, and thus readily accessible on the timescale of the simulations.

TIGHT-BINDING RELAXATION

In order to sit the above empirical models on physically more robust grounds, further relaxation was carried out using MD with tight-binding interatomic potentials, more specifically the Goodwin-Skinner-Pettifor (GSP) model for Si [8], and the Molteni-Colombo-Miglio (MCM) model for GaAs [9]. TB models are to some extent empirical (weighting of the matrix element as a function of distance, form of the repulsive interactions, cutoff radii, etc.), and thus have their limitations, but are nevertheless essential in the present context: First, there exists, to our knowledge, no empirical potential that can provide a reliable description of a-GaAs (we could not obtain satisfactory structures using the Tersoff potential [11]). Second, first-principles simulations are hardly feasible on such large systems as those considered here, on the timescale needed to achieve full relaxation. The TB method is an

Network	Si	GaAs	GaAs (SL)
CRN-P	-13.172	-13.450	
CRN-CT	-13.163	-13.561	-13.450
Crystal	-13.389	-13.802	-13.802

Table I: Energy (eV/atom) of the two networks relaxed with TB potentials at 0 K. For GaAs, we also give the results from the TB-MD simulations of Seong and Lewis (SL), Ref. [12].

excellent compromise between these two extremes, as it offers the transferability needed to describe accurately the various atomic environments encountered in disordered materials, in particular subtle details of the nearest-and second-nearest neighbour distributions, while, at the same time, providing information on the atomic structure of the material.

The models discussed in Section 2 assume the particles to be either all of the same type (CRN-P) or symmetric (CRN-CT). It is the purpose of the present article to assess the relevance of each model for a-Si and a-GaAs. We therefore relaxed *each* of the two models with *both* the GSP-TB model for Si and the MCM-TB model for GaAs, so that we have, in total, four different models: CRN-P-Si, CRN-P-GaAs, CRN-CT-Si, and CRN-CT-GaAs. In the case of CRN-P-GaAs, prior to relaxing, the sites of the network were labeled in such a way as to minimise the number of wrong bonds, leaving about 14% of them, close to the theoretical limit of about 12% [3].

RESULTS

Energetics

The four models were first brought to their ground state through static relaxation at 0 K. Surprisingly, it is found from Table I that models CRN-P-Si and CRN-CT-Si have essentially the same configurational energy, i.e., the elastic cost of eliminating odd-membered rings is very small, if present at all. However, there are many more ways of constructing a Polk network than a constrained Connell-Temkin network, and this is why the latter does not come out "naturally" from computer models. Also, this result indicates that there is really no competition between elastic strain and chemical ordering and that both can be fully accommodated by a CT-CRN. Indeed, the chemically-ordered CRN-CT-GaAs model is favoured over the CRN-P-GaAs by about 0.11 eV/atom (cf. Table I).

We have compared the energy of CRN-CT-GaAs with that of a 64-atom model obtained using the conventional melt-and-quench MD approach and the *same* set of TB parameters [12]. As can be seen in Table I, CRN-CT-GaAs lies a good 0.1 eV/atom in energy *below* the melt-and-quench model, indicating that the ART approach leads to a much better model than the usual MD thermal cycle; we will see more evidence for this below. In fact, for 216-atom models such as those considered here, detailed TB-MD simulations would hardly be feasible on the timescale needed to reach full convergence.

Following the static relaxation phase, the models were submitted to a dynamic annealing procedure in order to check for stability. We started with a 7.0 ps run at 300 K, followed by an 8.8 ps run at 800 K, after which the temperature was taken back at 300 K for another 3.5 ps, then at 10 K for a final 0.9 ps. This thermal cycle reveals the inadequacy of the Polk model for describing a-GaAs: Strong deformations of the CRN-P-GaAs model were observed at 800 K and the final structure lies 0.06 eV/atom *above* the initial one at the end of the cycle. Clearly, the system is not stable with respect to CRN-P, and got trapped (owing to the finite simulation time) into a high-energy state on its way to the ground state, which

CRN-P   **CRN-CT**
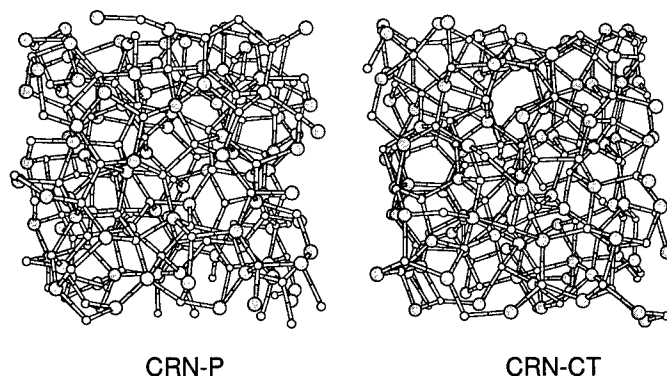
Figure 1: Ball-and-stick representation of models CRN-P and CRN-CT for GaAs; small white circles are As and larger grey circles are Ga atoms.

we argue is the CRN-CT-GaAs model. The latter, indeed, remained essentially unchanged during the annealing procedure, indicating that the structure lies in a minimum significantly deeper than CRN-P-GaAs. In view of these results, the discussion that follows pertains, for each model, to the initial 7.0-ps relaxation run at 300 K.

The two models for GaAs were analysed in detail in order to identify the effects of chemical (dis)ordering on the structural, electronic and dynamical properties and to assess them against available experimental data. Thus, in the remaining of this paper, we are concerned only with GaAs, and the models will be simply referred to as CRN-P and CRN-CT. A ball-and-stick representation of the two models is given in Fig. 1.

Structural properties

We compare in Table II the structural properties of models CRN-P and CRN-CT — both after the initial static (0 K) optimisation and after the 300 K relaxation run — with models obtained using either ab initio or TB MD. Since As and Ga atoms are treated in a symmetric manner during the ART optimisation, the local environments of the two types of atoms remains very similar even after the TB relaxation. This, in fact, is also observed in MD simulations of the material.

Table II shows that both CRN-P and CRN-CT are close to *ideal* realisations of the CRN, i.e. almost perfectly fourfold coordinated and a very narrow bond-angle distribution. These quantities compare favourably to those from the quenched samples. Moreover, CRN-CT is found to possess less than 4% of wrong bonds at 0K, and about 5% at 300 K, much less than the 14% of model CRN-P. Thus, CRN-CT probably constitutes the closest realisation of a CT-type CRN with periodic boundary conditions ever built. It should be noted that at 300 K, thermal agitation makes the identification of nearest-neighbours (and defects) difficult; we define them by using, as cutoff distance, the minimum between the first and second-neighbour peaks, indicated in Table II.

Recently, an EXAFS (extended x-ray absorption fine structure) study of a-GaAs produced by ion implantation was reported [15]. The average nearest-neighbour distance was found to be 2.451 ± 0.005 Å, a bit larger than in the crystal (2.442 Å), in agreement with density measurements [16]. This is a bit smaller than our own value of 2.51 Å and probably reflects some limitations of the model (incomplete relaxation, transferability of the TB

| | CRN-P | | CRN-CT | | SL | MCM | CP |
	0 K	300 K	0 K	300 K	0 K	0 K	0 K
$Z = 3$	0.046	0.128	0.051	0.118	0.242	0.14	0.219
$Z = 4$	0.954	0.845	0.944	0.830	0.598	0.66	0.781
$Z = 5$	0	0.026	0.005	0.045	0.129	0.18	0
$Z = 6$	0	0.001	0	0.004	0.024		0
$Z = 7$	0	0.000	0	0.002	0.007		0
$< Z >$	3.95	3.90	3.95	3.95	3.94	4.09	3.83
r_{NN} (Å)	3.0	3.1	3.0	3.1	3.0	3.0	2.8
r_1 (Å)	2.508	2.505	2.517	2.507			
Δr_1 (Å)	0.075	0.117	0.073	0.103			
Wrong bonds (%)	14.1	14.2	3.9	5.2	12.2	12.9	10.0
$\Delta\theta$ (deg.)	11.0	14.1	10.8	15.0	17.0	17.0	

Table II: Structural characteristics of the two models for GaAs discussed in the text, at both 0 and 300 K: distribution of coordination numbers, Z (and nearest-neighbour cutoff distance, r_{NN}), first nearest-neighbour distance, r_1 (and width, Δr_1), density of wrong bonds, and width of the bond-angle distribution, $\Delta\theta$. Also given are the corresponding numbers from other simulations, all at 0 K: SL — TB simulations of Ref. [12]; MCM — TB simulations of Ref. [13]; CP — Car-Parrinello simulations of Ref. [14].

model, etc.). For the average (total) coordination number, Ridgway et al. find 3.85 ± 0.2. This is certainly consistent with our own value of 3.95. However, coordination numbers are notably difficult to extract from EXAFS measurements; it would be of interest that x-ray measurements be carried out on ion-implanted a-GaAs, as has been the case for a-Si [17].

We show in Fig. 2 the total and partial structure factors for both models. The partials are not available from experiment, but the total is. We find the CRN-CT model to be in slightly better agreement with the data than CRN-P, but the differences are small and probably not very meaningful. Further, the experimental samples suffer from inhomogeneities and more accurate data are probably needed for a meaningful comparison to be made [18]. Likewise, the partial structure factors differ in many ways, but no signature of the presence of wrong bonds can be found in CRN-P: there are no feature present in CRN-P that does not show up also in CRN-CT.

While the partial static structure factors do not reveal qualitative differences between the two networks, the partial radial distribution functions (RDF's) do provide a direct way of probing the local environments and in particular measuring the density of wrong bonds. The partial pair correlation functions are however not available from experiment because of the similarity between As and Ga. This is quite unfortunate because the differences between the two topological models get almost completely washed out when averaging into the total RDF as can be seen from Fig. 3.

The partial RDF's are also displayed in Fig. 3. It is clear that CRN-CT is chemically ordered (i.e., wrong bonds are rare): the Ga-As correlation exhibits a strong first-neighbour peak, but the second and fourth peaks are very weak. In contrast, the like-atom functions, Ga-Ga and As-As, have no nearest-neighbour peak, but exhibit strong second and fourth-neighbour peaks. Thus, chemical ordering filters out the shell structure of the material.

The large split peak in the range 3.5-7 Å in the Ga-As RDF contains information about the distribution of dihedral angles. From it we can infer that dihedral angles of 60 and 180 degrees are present in significant numbers in both structures. This can actually be seen from a direct calculation of the distribution of dihedral angles, shown in Fig. 3 together with the
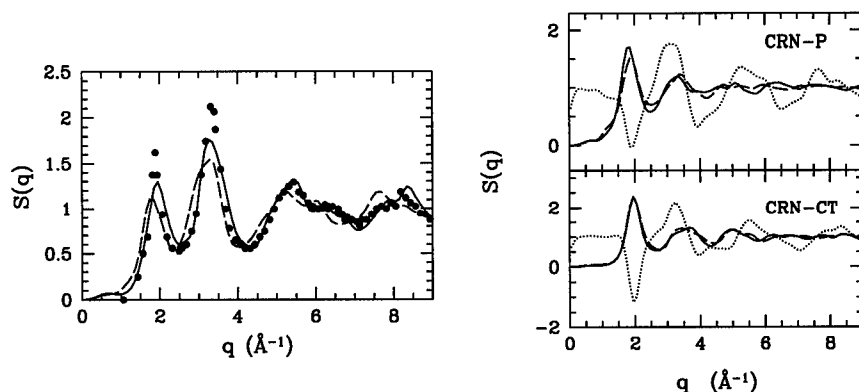
Figure 2: *Left:* Total static structure factors for CRN-P (dashed line) and CRN-CT (solid line); the dots are the experimental data of Udron *et al.*, Ref. [18]. *Right:* Partial static structure factors for the two models, as indicated; the dotted, dashed, and solid lines are for the Ga–Ga, Ga–As, and As–As partial correlations.

distribution of bond angles. The width of the bond-angle distribution, about 11 degrees at 0 K, is somewhat smaller than that obtained in previous models (cf. Table II) agrees quite closely with experiment, and is comparable with that measured in fully-optimised Wooten-Winer-Weaire models for Si [19]. The distributions of bond and dihedral angles are thus almost the same in the two models. This disagrees somewhat with the results of Connell and Temkin, who's mechanical model was found to differ significantly from Polk's model in this regard and can probably be attributed to the presence of surfaces in the mechanical models.

Electronic and vibrational properties

Although chemical ordering can be said to be essentially structural, it should manifest itself also on the electronic and vibrational properties of the material. X-ray photoemission-spectroscopy (XPS) measurements on a-GaAs have been interpreted as indicating almost complete chemical order [20]. Though the present study confirms this conclusion, as discussed above, it also suggests that the density of electron states (EDOS) is *not* a good probe of the presence of wrong bonds: As can be seen in Fig. 4, the EDOS for the two models are almost completely identical. (The three valence bands can be ascribed roughly as follows: A is composed of As and Ga p states, B arises from Ga s and some As p states, and C corresponds to the As s states [21, 22]). Although the ionicity gap (at -9 eV) and the conductivity gap are slightly larger for CRN-CT than for CRN-P, the general features of the two EDOS are similar enough that they probably could not be distinguished by an XPS experiment.

A similar situation is found with the vibrational properties. The vibrational density of states (VDOS) for both models CRN-P and CRN-CT, displayed in Fig. 4, give similar agreement with the VDOS extracted from Raman spectroscopic measurements [23]. The TA and TO peaks are roughly at their right position but the respective weights of the two peaks differ: experimentally, the weight of the TA peak is slightly smaller than that of the TO peak. This would indicate that the number of coordination defects in both CRN-P and CRN-CT still remains greater than that of the real material [24, 25]. Such conclusion is

469
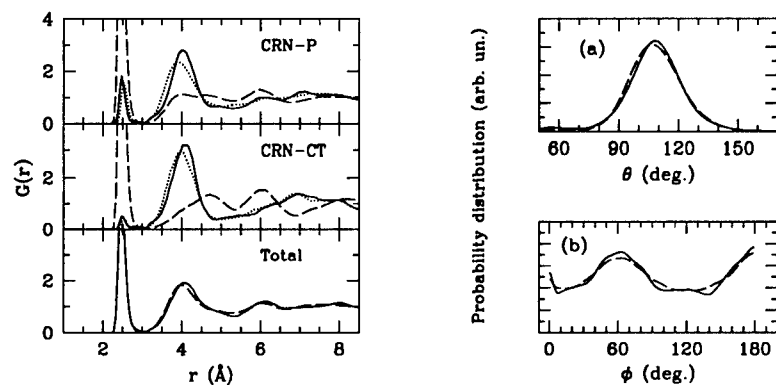
Figure 3: *Left:* partial RDF's for the two models, as indicated; the dotted, dashed, and solid lines are for the Ga–Ga, Ga–As, and As–As partial correlations, respectively. The lower panel gives the total (unweighted) RDF for CRN-P (dashed line) and CRN-CT (solid line). *Right:* Distributions of (a) bond and (b) dihedral angles for CRN-P (dashed lines) and CRN-CT (full lines).
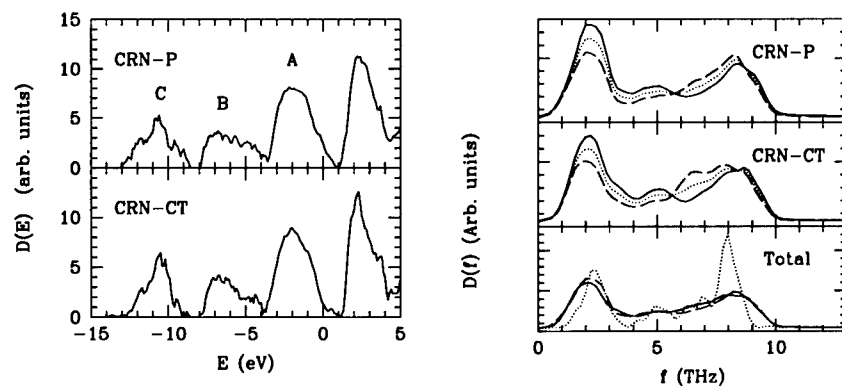
Figure 4: *Left:* Electronic densities of states for the two models, as indicated. The identification of the peaks is discussed in the text. *Right:* Partial and total vibrational densities of states for the two models, as indicated. Dashed lines are for Ga atoms, solid lines for As and dotted lines are the totals. The lower panel presents a comparison of CRN-P (dashed line) and CRN-CT (solid line) with the crystal (dotted line).

also supported by the fact that the TO peak calculated from a TB-MD/melt-and-quench simulation of *a*-GaAs, which contains a higher density of defects than our models (cf. Table II), is almost nonexistent. The coordination of *a*-GaAs must therefore be almost exactly 4, with only a very small density of defects.

CONCLUSION

Using the activation-relaxation technique of Barkema and Mousseau in conjunction with empirical potentials, we have constructed almost-perfect model structures for continuous random networks exhibiting a significantly different intermediate-range character. Because of the generic approach used here, the physical relevance of the models is not determined *a priori* but, rather, *a posteriori* through detailed relaxation under realistic tight-binding potentials. This, further, has enabled us to examine the energetics and stability of the networks. We have found, in this way, that a Polk-type model for GaAs is not stable even under a modest heat treatment, while a Connell-Temkin model can support annealing (on the MD timescale) without going through structural changes. Though we cannot claim to have found *the* ultimate structural model for *a*-GaAs, these observations, together with the fact that both models are energetically equivalent in the case of Si, strongly support our conclusions.

This mixed-approach used here have allowed us to compare the properties of two different models for *a*-GaAs in the hope of finding a possible signature of the presence of wrong bonds. It is unfortunate that very few quantities are sensitive to the presence of wrong bonds, even at such levels as 5—15%. Electronic and vibrational densities of states are essentially undistinguishable for the two networks. It appears that only direct measurements of partial correlation functions by diffraction can provide the information needed to establish the density of wrong bonds in *a*-GaAs and, therefore, the nature of the intermediate-range topology of amorphous semiconductors.

For many optimisation problems in materials science, a single approach is not practical. We have shown here that a combination of techniques and potentials can provide significant insights into some problems. In particular, it was possible to limit ourselves to relatively short TB-MD runs in order to establish the stability of our models and thus achieve a state of convergence which would be prohibitively expensive through direct TB or *ab initio* simulations.

REFERENCES

[a] E-mail: lewis@physcn.umontreal.ca.

[b] Permanent address: Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA; e-mail: mousseau@helios.phy.ohiou.edu.

1. J. J. Dong, D. A. Drabold, Phys. Rev. B **54**, 10284 (1996).

2. D.E. Polk, J. Non-Cryst. Sol. **5**, 365 (1971).

3. M.-L. Thèye, A. Gheorghiu, and H. Launois, J. Phys. C.: Solid St. Phys. **13**, 6569 (1980).

4. G.A.N. Connell and R.J. Temkin, Phys. Rev. B **9**, 5323 (1974).

5. J. C. Phillips, *Bonds and Bands in Semiconductors* (Academic, New York, 1973).

6. G.T. Barkema and N. Mousseau, Phys. Rev. Lett. **77**, 4358 (1996).

7. N. Mousseau and G. T. Barkema, Phys. Rev. E **57**, to be published.

8. L. Goodwin, A.J. Skinner, and D.G. Pettifor, Europhys. Lett. **9**, 701 (1989).

9. C. Molteni, L. Colombo, and L. Miglio, J. Phys.: Condens. Matter **6**, 5243 (1994).

10. F.H. Stillinger and T.A. Weber, Phys. Rev. B **31**, 5262 (1985).

11. J. Tersoff, Phys. Rev. B **39**, 5566 (1989); R. Smith, Nucl. Instr. Meth. in Phys. Res. B **67**, 335 (1992); M. Sayed, J.H. Jefferson, A.B. Walker, and A.G. Gullis, Nucl. Instr. Meth. in Phys. Res. B **102**, 232 (1995).

12. H. Seong and L.J. Lewis, Phys. Rev. B **53**, 4408 (1996).

13. C. Molteni, L. Colombo, and L. Miglio, Phys. Rev. B **50**, 4371 (1994).

14. E. Fois, A. Selloni, G. Pastore, Q.-M. Zhang, and R. Car, Phys. Rev. B **45**, 13 378 (1992).

15. M.C. Ridgway, C.J. Glover, G.J. Foran, and K.M. Yu, preprint.

16. L. Cliche, S. Roorda, and R.A. Masut, Appl. Phys. Lett. **65**, 1754 (1994).

17. K. Laaziri and S. Roorda, private communication.

18. D. Udron, M.-L. Thèye, D. Raoux, A.-M. Flank, P. Lagarde, and J.-P. Gaspard, J. Non-Cryst. Solids **137&138**, 131 (1991).

19. F. Wooten and D. Weaire, Solid State Physics **40**, 1 (1987).

20. C. Sénémaud, E. Belin, A. Gheorghiu, and M.-L. Thèye, Solid St. Comm. **55**, 947 (1985).

21. J.D. Joannopoulos and M.L. Cohen, Phys. Rev. B **10**, 1545 (1974).

22. E.P. O'reilly and J. Robertson, Phys. Rev. B **34**, 8684 (1986).

23. A. Chehaidar, A. Zwick, R. Carles, and J. Bandet, Phys. Rev. B **50**, 5345 (1994).

24. N. Mousseau and L.J. Lewis, Phys. Rev. B **43**, 9810 (1991).

25. R. Alben, D. Weaire, J.E. Smith, and M.H. Brodsky, Phys. Rev. B **11**, 2271 (1975).

A NOVEL SCHEME FOR ACCURATE MD SIMULATIONS OF LARGE SYSTEMS

Alessandro De Vita† and Roberto Car
Institut Romand de Recherche Numérique en Physique des Matériaux (IRRMA), PPH-Ecublens, CH-1015 Lausanne, Switzerland

ABSTRACT

We present a simple and informationally efficient approach to electronic-structure-based simulations of large material science systems. The algorithm is based on a flexible embedding scheme, in which the parameters of a model potential are fitted at run time to some precise information relevant to localised portions of the system. Such information is computed separately on small subsystems by electronic-structure "black box" subprograms, e.g. based on tight-binding and/or ab initio models. The scheme allows to enforce electronic structure precision only when and where needed, and to minimise the computed information within a desired accuracy, which can be systematically controlled. Moreover, it is inherently linear scaling, and highly suitable for modern parallel platforms, including those based on non-uniform processing. The method is demonstrated by performing computations of tight-binding accuracy on solid state systems in the ten thousand atoms size scale.

INTRODUCTION

Modelling materials at the atomic level via Molecular Dynamics (MD) techniques is often a computationally intensive task. In spite of the ever-increasing availability of computer power, all known MD techniques are bound to be limited. The limits concern *e.g.*, the size of the systems studied, the physical times simulated, and the precision of the force-models used. The exact balance of the simulation parameters depends on the nature of the specific application, and is often a compromise between the competing requirements of large enough simulated sizes, sufficient accuracy, and sufficient amount of accumulated statistics. Broadly speaking, the most accurate (*e.g.* first principles) simulations are limited to systems of a few hundred atoms and to simulation times in the picosecond range. Simulations using classical model potentials can currently deal with several million atoms (see *e.g.*, ref. [1]) and with the nanosecond time range.

Our discussion of the current status of MD techniques starts by looking at how the computed information is normally used to determine the simulated trajectories. It turns out that for a broad range of high-accuracy applications, once a concept of a non-uniform tolerable error is specified, it is possible to speed up the computations by a non-negligible factor. This can be achieved by saving on the amount of redundant computed information. Redundancies occur if the MD algorihm used is not flexible enough to evaluate new information only *wherever* and *whenever* this is really needed.

An appealing scheme to introduce such flexibility can be developed from the concept of *teaching on the flight* to a classical model potential *about* all the "difficult" situations occurring anywhere in the system during the simulation run. The "difficult" situations correspond in practice to those local atomic arrangements for which the model alone is not a priori sufficiently accurate. For them, a more precise (and "costly") technique is used to develop support information. In this way, all the necessary information is fed into the calculation, while computer intensive tasks are performed only where and when this is really needed.

A central feature of the scheme is that these computer intensive tasks need not to identify with calculations on the entire (and, may be, very large) original system, or on exact subsets of it. Rather, they are calculations on small systems whose only requirement is to be *relevant* to the dynamical history of the original system, but which can otherwise be designed for computational convenience. A possible procedure will be briefly discussed in the next sections, describing how these features are included into a practical MD implementation. A detailed account of the method will appear separately.[2] Here we only note that, from an abstract point of view, lowering the amount of computed information to the minimum compatible with the required precision is *the only* way to achieve a given description of a physical system (the atomic trajectories in MD) in a shorter time at fixed computing power.

HANDLING OF ACCURACY IN MD SCHEMES

Most MD techniques adopt by construction a uniform accuracy in space and time: the same force routines are called at every time step, throughout the whole system[1]. A known way to improve on this situation consist of eliminating from the calculation some frequency-decoupled degrees of freedom, which can be neglected as irrelevant to the physical problem under study. This is usually done by using constraints, as *e.g.* in those classical simulations on water where the geometry of the molecule is kept fixed. If a subset of the forces in the system is much larger than the rest -but constraining the "fast" variables is a too crude an approximation- some multiple time scale methods exist which integrate separately, with different time steps, the slow and fast degrees of freedom. This may be done *e.g.* by exploiting an appropriately chosen analytically solvable reference system for the high frequency motion (cfr. ref. [3] and references therein).

A different class of MD applications are those in which the focus is on some specific part of the system, usually the "central" one, whose surrounding regions are of less importance, and typically are represented only to provide the correct (chemical, dielectric, elastic) termination. The atoms in the central region are moved according to forces calculated with some high precision scheme (*e.g.* an ab initio scheme) while the surrounding region is described by an appropriate, less sophisticated model. In this class of schemes (known as "embedding" schemes) the computations at different levels of accuracy are, in other words, performed *directly on appropriate subportions of the system studied.* The success of a particular embedding strategy is achieved by making an appropriate choice of the model for the outer region and by overcoming the difficulties of matching different schemes on the interface region.[2]

There is, however, no immediate reason why achieving a high level of accuracy on a portion of a physical system should imply matching problems. Suppose for example that a force model capable of "learning" from a high-quality ("exact") source of information (say, a "black box ab initio engine") were fed data concerning preeminently a particular portion of the system. Taking those data into account, the model would then produce a time evolution which would be particularly precise in that specific portion of the system where precision is sought for (but not necessarily elsewhere). The needed "importance sampling" of precision would be automatically achieved. The difficulties of matching different models on border regions would be removed as long as the "learning" procedure were sufficiently smooth in merging apprehended information with previous knowledge.

---

[1]In the special case of first principles simulations based on the momentum-space formalism, the uniform accuracy in representing the electronic degrees of freedom is extended to the whole simulation box, regardless if the system only occupies a part of it

[2]A very interesting approach is the recently introduced Quasicontinuum Method, aimed at the study of multiple-scale phenomena in crystalline solids. The method offers a way to merge the atomistic description with a finite element formalism, thus allowing for the high level of accuracy of the Newtonian dynamics in chosen "refined" regions[4]. The transfer of heat and linear momentum from fully refined regions to the far-field regions can be performed by *e.g.* constrained dynamics[5], while the merging can be *analytically* described by appropriate hamiltonian elastic formalisms[6].

Moreover, highly expensive information could be fed into the model in larger proportion during the specifically interesting steps of the simulated process, and much less before or after. This would be a straightforward way to perform importance sampling of computed information throughout the *time* variable. We can refer to this as "time embedding", as opposed to the "space embedding" considered above.[3] From what seen to this point, the use of a force model capable of learning may allow to compute "exact" information only if necessary: only if the model cannot guarantee high precision when and where such precision is really needed by the scope of the calculation.

CLASSICAL MODELS, PRECISION, AND "ON FLIGHT" TEACHING

For the purpose of this section we can define as our "exact" scheme any arbitrary computational scheme which takes explicitly into account the electronic degrees of freedom of the system. (The scheme may be based, *e.g.* on a tight-binding or on a first principles single particle formulation.) Computing the inter-atomic forces with a technique of this kind is inherently much slower than using a classical model, where the forces are obtained analytically from a parametrised potential form $V^P(\{\mathbf{R}_I\}, \{\alpha_i\})$.

A classical potential able to reproduce reliably the "exact" forces for any given atomic configuration of the system would clearly be an ideal MD tool. Indeed, classical potentials are often constructed so to reproduce "exact" computed data (and/or experimental data). However, a vast body of experience shows that not all materials are easy to model with classical potentials and that for a given material a normally well-behaved potential may well fail miserably in particular circumstances (indicating a problem of *transferability*). This is most likely to happen when the system visits a configuration very different from those contained in the model fitting data base, or when the analytic form of the potential simply cannot capture (for any fixed choice of the potential parameters) the essence of the physical processes taking place. Transition states, bond breaking, and hybridisation changes in covalent materials, localisation and delocalisation of electrons in π-bonded structures are typical "difficult" situations in which classical potentials may be expected to fail.

Obtaining reliable inter-atomic forces with a given parametrised potential form and an appropriately fixed set of parameters can thus be a difficult task. The main problem is that although most potentials are known to be very good in some region of the system's phase space, the extension of this region of validity is never perfectly well known[7], and very hard to control. This is a crucial issue when fitting model potentials: if the applications require that the validity region be considerably wide, this poses extreme demands on the potential form. A safe procedure during a MD run would be to check the model versus the exact forces each time the system passes through a critical point of the simulated process, *i.e.* one in which the potential may a priori fail. The whole MD trajectory would have to be rejected if any of such tests failed [4].

This kind of "on flight" testing is very rarely done, and for large system sizes it would be, anyway, unfeasible. Even when feasible, testing with the exact model would be likely to take (much) more time than the classical MD run itself. Still, the procedure suggests a further development in which, instead of rejecting the trajectories, the exact forces evaluated are introduced as new knowledge into the classical potential. To do this, it is necessary to release the constraint of a *fixed* set of parameters $\{\alpha_i\}$ in the model potential $V^P(\{\mathbf{R}_I\}, \{\alpha_i\})$. Indeed, the parameters can be made to vary along the trajectory, and become a function of time $\{\alpha_i(t)\}$. This can be done *e.g.* by a fitting procedure in parameter space at different times. The optimal set of parametric

---

[3]We have assumed here that the potential model can learn and correctly reproduce information which is "local" in space and/or in time. This is not too difficult to achieve *e.g.*, by using time dependent "local" parameters in the model, as explained in the next sections.

[4]A more complete analysis of these problems, here only summarised in the text, can be found in ref.[7]

functions $\{\alpha_i(t)\}$ is in this case that which best reproduces the exact forces at all times along the atomic trajectory $\mathbf{R_I(t)}$.

The scheme can be summarised as follows: (i) during a classical simulation using the potential $V^P(\{\mathbf{R_I}\}, \{\alpha_i\})$, "difficult" subsystems are identified in the whole system studied (the "difficult" event could be *e.g.*, a chemical bond breaking at some time t_0; the subsystem would then contain the atoms located in a region immediately surrounding the bond-breaking site); (ii) a new small system is constructed for each of these subsystems, incorporating the subsystem's relevant features but otherwise designed to be treated efficiently by an exact computation. This small system may be constructed in various ways (it could be a *cluster* centred on the breaking bond and composed of its nearest neighbouring atoms taken at the positions they occupy at time t_0. The cluster could be chemically terminated, or a few more atoms could be added to it, to give a small periodic system[5]). An exact computation is at this point performed on each small system.[6] (iii) The exact results are "taught" to the model. (Using a short-range potential form, this will mostly influence the values of those parameter functions $\{\alpha_i(t)\}$ which are relevant to the region surrounding the breaking bond, for $t \sim t_0$). The model should now be able to tackle successfully the difficult configurations encountered in the original MD simulation.

In this scheme the exact computations can be expected to take most of the computer time. Although some overhead is implicit in point (ii) above (*e.g.* if an exact calculation on a full cluster is performed to help representing the breaking of a bond between *two* atoms), the flexibility and, as discussed, the informational efficiency of the method should more than compensate for this in several situations of interest. Moreover, accuracy can be tested and tuned in any circumstance (*e.g.*, by increasing the cluster sizes), and exact models are trivially easy to include, substitute, or *combine* as desired in a given application. Also, the computations required by the scheme scale linearly with the size of the system studied. Finally, we note that even a simple parametrised potential form can be expected to provide accurate forces if the parameters are adjusted "on the flight" for the scope. With the enhanced "representation" power due to parameter varying, a straightforward potential expansion up to the three-body term can *e.g.* represent successfully the forces of a system as complicated as a covalent liquid (see next section). As noticed above, keeping the parameters fixed may simply *never* guarantee very high accuracy for such a system, even using fairly complicated analytical form. It is only in some cases that an enormous range of configuration space can be well described by a simple model with fixed parameters, as in the case of the 'glue' model potentials for metallic systems[9], where inter-atomic potentials of very good quality can be constructed by *e.g.* force-matching procedures on very large data bases of first-principles data[10].

A SIMPLE APPLICATION

The frequency with which exact computations have to be performed depends on the initial quality of the model potential and on the history of the studied system. However, once the model has been chosen and a tolerable error has been mathematically specified, it is only a matter of computing enough 'exact' information, and transmitting it efficiently to the model. As an example, we report here some result on liquid Si obtained by a practical implementation of

---

[5]This second option may be convenient to optimise size-convergence, and especially so if the exact technique adopted works optimally in periodic boundary conditions, as is the case for standard plane-wave first principles implementations.

[6]Here we make the crucial assumption that an exact quantum mechanical calculation on a *small* system can be relevant to the local properties of the original extended system. There may be difficulties with this (*e.g.*, if the original system is a metal at T=0, even if the "small" system is periodic), but the problem can be handled *e.g.*, by using finite electronic temperature schemes. A discussion of a "nearsightedness" principle underlying these considerations can be found in reference[8].

the approach so far described. A full account of these results and a detailed description of the implementation will be included in a more systematic technical report[2].

Liquid Si is a prototypical case in which the results of a method based on the explicit description of the electronic wave functions are difficult to reproduce with classical potentials. We choose as our 'exact' model the tight-binding model of ref.[11] and as classical model the Stillinger-Weber (SW) potential (containing two- and three-body terms) with initial parameters as in ref.[12]. Since the system has a uniform behaviour in both space and time, so that no "embedding" strategy can be exploited to save on computations, we choose here to perform an 'exact' computation on *each* atom in the system every $N_s \sim 10$ dynamical steps. This is done by computing the force on the atom when surrounded by a cluster of its neighbours "carved" out of the original system and chemically terminated with H atoms (the Si-H tight binding model was taken from ref.[13]). A force matching fit of the SW parameters is then performed to determine the parameters $\{\alpha_i^n\}$ for which the model forces $\mathbf{F_J}^n = -\partial V^P(\{\mathbf{R_I}\}, \{\alpha_i^n\})/\partial \mathbf{R_J}$ match the forces computed exactly by the cluster calculation at time step n. [7] Note that the latter constitute only incomplete information about the true forces in the original system. For example, the "exact" forces computed using clusters will not in general add up to exactly zero, while the fitted ones $\mathbf{F_J}^n$ will (for an isolated system). The model potential is used at this point to generate N_s more MD steps before a new fit is performed (a time step of 10 a.u., and a Nose thermostat were used to simulate liquid Si at 2000K).

Notice that to each set of parameters in the sequence $\{\alpha_i^j\}, j = n, n + N_s, \ldots$ corresponds a domain of high accuracy in the space of atomic coordinates. We can think this domain as centred on the configuration space point $\{\mathbf{R_I}(t_j)\}$ for which the set was fitted, and to include atomic configurations $\{\mathbf{R_I}'(t_j)\} \simeq \{\mathbf{R_I}(t_j)\}$ which are close but slightly off those of the fit. This can be exploited to improve the atomic trajectories by simply repeating the N_s MD steps preceding the latest fit (note that classical MD steps are computationally inexpensive). This time, the parametric function values $\{\alpha_i(t_j)\}$ which determine the repeated steps via the interatomic potential are generated by *interpolation* of fitted values. This gives enhanced accuracy and slightly corrected trajectories without introducing the need of additional fits. Using these recipes, preliminary tests indicate that an average relative error of $\sim 10\%$ or less is made on atomic force components larger than $\sim$0.5 eV/Å in liquid Si, if H-terminated clusters of about one hundred atoms or larger are used.

Figure 1 shows the pair distribution function $g(r)$ computed for a 216-atom system of liquid Si at bulk density at a temperature of 2000K (solid line). The dotted line corresponds to the results obtained with the exact model dynamics, *i.e.* by diagonalising the tight binding hamiltonian of the full system at each time step. The residual differences of the two curves are due to the errors introduced by the cluster scheme and time interpolation procedure. The dot-dashed curve corresponds to the $g(r)$ one would obtain by the SW potential alone, without fitting it to the tight binding model. Test on dynamical properties such as the velocity autocorrelation function show an agreement similar to the one in the figure.

---

[7]The α_i parameters of the two-body and three-body terms are allowed to differ for different couples and triplets of atoms in the system, so that the individual parameters acquire a "local" meaning in the short ranged SW form.
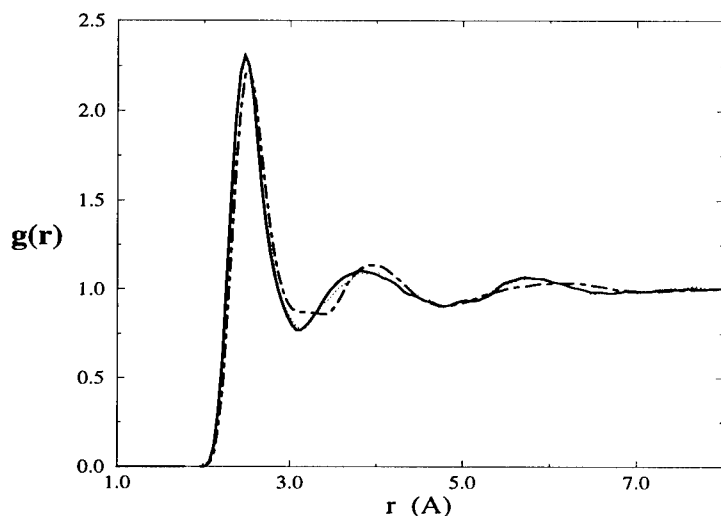
Figure 1. Pair correlation function in liquid Si (see text).

The MD approach presented in this work is aimed at the study of processes in which multiple scales may be involved, so that it may be necessary to perform computations on large systems. To investigate the feasibility of these computations, we coded a parallel implementation of the algorithm described above. We tested the code by performing calculations reproducing the tight-binding 'exact' results with the same accuracy as above on systems up to about 20000 atoms. A typical simulation geometry is illustrated in figure 2, showing a 11200-atom system of crystalline bulk Si in a tetragonal cell of about 220 nm^3 volume.[8] The figure also shows the plot of the system's temperature vs. time for a time span of about 0.4 ps after thermalisation at 500K.

The parallel coding of the algorithm proved particularly simple. Efficient linear scaling algorithms for implementing classical MD are readily available in textbooks (for the system sizes of interest here, setting up a classical MD application does not require extreme sophistication on the currently available machines). Virtually all the computer time is spent in exact computations

---

[8]For production runs on such a geometry we typically use eight SGI 195MHz RS10000 processors.

478

even using as "exact" model the relatively simple tight-binding model.
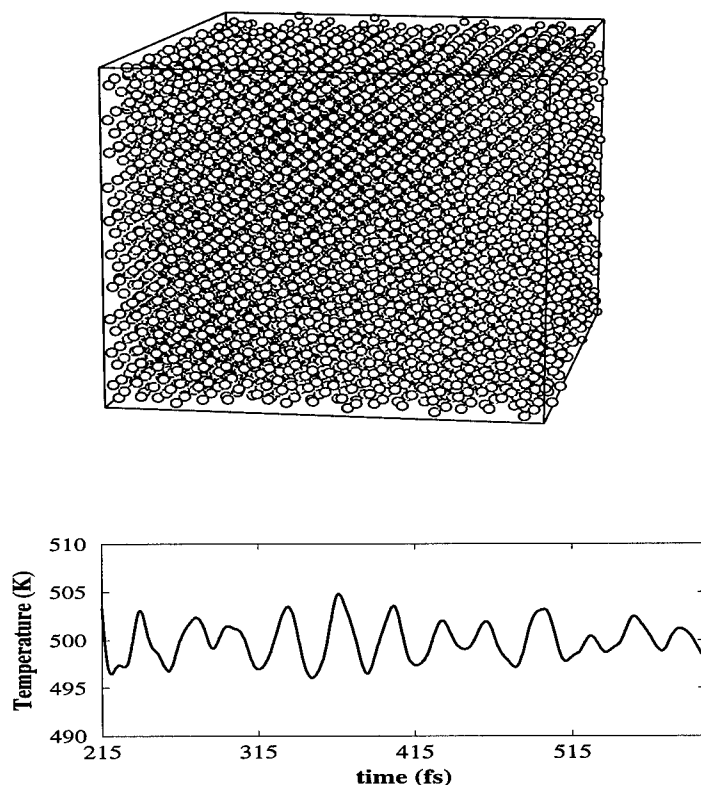
Figure 2. The 11200 atom test system, and temperature vs. time plot.

The exact computations are logically decoupled, and their individual core memory require-
ments do not add up to determine a total workload.[9] The data which need to be stored in mem-
ory, like the atomic positions and the fit parameters, are used for CPU inexpensive operations,
while the CPU-intensive and memory intensive information, related to the electronic degrees of
freedom, is developed as processor-private data, and never broadcasted. Thus the algorithm is,
in practice, "embarrassingly parallel" for its almost negligible requirements of communication
bandwidth, and very moderately requiring in terms of memory. This makes it suitable for pro-
duction on massively parallel computing platforms with high row power and comparatively low
communication performance, *and* on smaller parallel platforms equipped with a (even only mod-
erately performing) shared memory option. Finally, dynamical load balancing can be achieved

[9]An O(N) memory complexity may be introduced by developments of the algorithm requiring Fermi energy
alignment between different exact quantum mechanical computations. See the discussion in ref.[14] in the context
of a well established O(N) method whose "divide and conquer" spirit has some point of contact with the method
presented here.

by unevenly distributing across processors the "exact" computations, which makes the scheme efficient on platforms based on non-uniform processing (heterogeneous clusters of workstations, or departmental parallel machines whose processors work in time sharing).

---

† Also at DIMCA, via A. Valerio 2, I-34127, Trieste.

REFERENCES

1. S.J.Zhou,D.M.Beazley,P.S.Lomdahl and B.L.Holian, Phys. Rev. Lett. **78** 479, (1997).

2. A.De Vita and R.Car, in preparation.

3. M.E.Tuckermann, G.J.Martyna and B.J.Berne, J.Chem.Phys. **93**(2), 1287 (1990); M.E.Tuckermann and B.J.Berne, J.Chem.Phys. **94**(2), 1465 (1991).

4. Cfr. *e.g.* abstracts P 1.7 (E.B.Tadmor, E.Kaxiras and N.Bernstein), P 5.1 (V.B.Shenoy and R.Phillips), P 9.6 (D.E.Rodney and R.Phillips) and *P 9.8 (R.Phillips, W.W.Gerberich, D.Kramer and E.B.Tadmor) in the MRS 1997 Fall Meeting Abstracts Book, and related articles in this book of proceedings.

5. Cfr. abstract *P 9.8 (R.Phillips, W.W.Gerberich, D.Kramer and E.B.Tadmor) in the MRS 1997 Fall Meeting Abstracts Book.

6. J.Q.Broughton, private communication.

7. M.C.Payne, I.Stich, A.De Vita, M.J.Gillan and L.J.Clarke, Faraday Discuss. **96**, 151 (1993).

8. W.Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

9. F.Ercolessi, E.Tosatti and M.Parrinello, Phys. Rev. Lett. **57**,719 (1986); M.S.Daw and M.I.Baskes, Phys. Rev B **29**, 6443 (1984); M.W.Finnis and J.E.Sinclair, Phil. Mag A **50**, 45 (1984).

10. F.Ercolessi and J.B.Adams, Europhys. Lett. **26**, 583 (1994).

11. Goodwin et al., Europhys. Lett. 9, 701 (1989).

12. T.A.Weber and F.H.Stillinger, J. Chem. Phys. **92**, 6239 (1990).

13. G.Panzarini and L.Colombo, Phys.Rev.Lett. 73, 1636 (1994); G.Panzarini and L.Colombo, Phase Transitions 52, 137 (1994).

14. T.-S.Lee, D.M.York and W.Yang, J.Chem.Phys. **105**, 2744 (1996); W.Yang, Phys.Rev.Lett.**66**, 1438 (1991).

LARGE SCALE ATOMISTIC SIMULATIONS USING THE TIGHT BINDING APPROACH

M.CELINO[1], F.CLERI[2], L.COLOMBO[3], M.ROSATI[4], V.ROSATO[1], J.TILSON[5]
[1]ENEA, HPCN Project, C.R. "Casaccia", C.P. 2400, 00100 Roma A.D. (Italy)
[2]ENEA, Dip. Innovazione, C.R. "Casaccia", C.P. 2400, 00100 Roma A.D. (Italy)
[3]INFM and Dip. di Scienza dei Materiali, Univ. di Milano, via Emanueli 15, 20126 Milano (Italy)
[4]CASPUR, Università "La Sapienza", P.le A.Moro 5, 00100 Roma (Italy)
[5]Mathematics and Computers Division, Argonne National Lab., Argonne, IL 60439 (USA)

ABSTRACT

Atomistic modelling of Materials Science problems often requires the simulation of systems with an irreducibly-large unit cell, such as amorphous materials, fullerites, or systems containing extended defects, such as dislocations, cracks or grain boundaries. Large-scale simulations with the Tight-Binding approach must face the computational obstacle represented by the $O(N^3)$-scaling of the diagonalization of the Hamiltonian matrix. This bottleneck can be overcome by parallel computing techniques and/or the introduction of faster, $O(N)$-scaling algorithms. We report the activities performed in the frame of a collaboration among several research groups on the porting of TBMD codes on parallel computers. In particular, we describe the porting of a $O(N^3)$ TBMD code on different MIMD computers, with either distributed or shared memory, by using appropriate software tools. Furthermore, preliminary results obtained in the porting of an $O(N)$ TBMD code on an experimental, hybrid MIMD-SIMD computer architecture are reported. The new perspective of using specialized platforms to deal with large-scale TBMD simulation is discussed.

INTRODUCTION

The Tight-Binding Molecular-Dynamics (TBMD) approach is becoming widespread in the atomistic simulation community. The success of TBMD stands on a good balance between the accuracy of the physical representation of the atomic interactions and the resulting computational cost. TBMD implements an empirical parametrization of the bonding interactions based on the expansion of the electronic wave functions on a very simple basis set. Thus, contrarily to classical Molecular Dynamics (MD) [1], TBMD allows to evaluate both ionic and electronic properties.

Although being much simpler than so-called *ab-initio* approaches [2], the computational complexity of TBMD algorithms is still considerable. The main limitation arises from the $O(N^3)$ scaling of the diagonalization of the Hamiltonian matrix at each time step. As a consequence, the practical size of the simulated systems cannot exceed the limit of 3-400 atoms on a workstation. While such a figure already allows the study of a number of relevant problems in semiconductor physics, it still leaves a large gap towards the true domain of Materials Science simulations, *e.g.* grain boundaries, dislocations, interfaces, nanostructures and so on.

This computational bottleneck can be overcome by either (a) the introduction of more efficient algorithms, having a more convenient $O(N^2)$ or $O(N)$ scaling, or (b) the use of parallel platforms to perform the diagonalization procedure. In this way, the study of systems in the range of about 10^3 atoms could become accessible. Furthermore, coupling of efficient $O(N)$ algorithms with parallelization techniques would allow to reach even larger size limits, in the range of 10^4 atoms.

In this paper we report two examples of porting of $O(N^3)$ [3] and $O(N)$ [4] TBMD codes on parallel computers with very different architectures. In particular, the $O(N^3)$ code was ported on two MIMD (Multiple Instruction Multiple Data) parallel computers using in each case different languages and mathematical libraries. The $O(N)$ code, in turn, was ported on a hybrid MIMD-SIMD (Single Instruction Multiple Data) parallel platform with the aim of assessing the usability of this new kind of high-performance computer architecture in large-scale simulations.

The paper will be divided into two parts: in the first one, the $O(N^3)$ and the $O(N)$ TBMD codes will be analyzed from an algorithmic point of view, enlightening the main issues involved in the parallelization. In the second part, the implementation of the two codes on the different architectures will be discussed.

$O(N^3)$ AND $O(N)$ TBMD: FROM THE SERIAL TO THE PARALLEL CODE

The main difficulties in the parallelization of such codes are: (1) the system cannot be mapped on a regular grid (atoms can move inside the simulation box); (2) in the case of the $O(N^3)$ code a diagonalization of a large sparse matrix is involved (parallel math libraries for sparse matrix diagonalization are not yet available); (3) in the case of the $O(N)$ code on the particular SIMD computers which has been considered, problems arise from the lack of local integer addressing on the processing nodes.

$O(N^3)$ TBMD code

In a way completely similar to a classical MD code, the $O(N^3)$ TBMD code can study atomistic systems in different thermodynamic ensembles when coupled to the Parrinello-Rahman-Nosé (PRN) boundary conditions [5].

In the adiabatic approximation the Hamiltonian of a system of atoms and electrons in a solid can be written as [3]

$$H_{tot} = T_i + T_e + U_{ee} + U_{ei} + U_{ii} \tag{1}$$

where $T_{i,e}$ is the kinetic energy of ions and electrons, U_{ee}, U_{ei}, U_{ii} are the electron-electron, electron-ion and ion-ion interactions, respectively.

Referring to the theory of one electron moving in the presence of the average field due to the other valence electrons and ions, the reduced one-electron Hamiltonian can be written $h = T + U_{ee} + U_{ei}$, giving the eigenvalues (energy levels) ϵ_n and the eigenfunctions $|\Psi_n>$. In a TB scheme, the eigenfunctions are represented as a linear combination of atomic orbitals $|\phi_{ln}>$

$$|\Psi_n> = \sum_{l\alpha} c_{l\alpha}^n |\phi_{l\alpha}> \tag{2}$$

where l is the quantum number index and α labels the ions. The expansion coefficients $c_{l\alpha}^n$ represent the occupancy of the l-th orbital located at the α-th site.

In the present TB approach, the elements of the h matrix, $<\phi_{l'\beta}|h|\phi_{l\alpha}>$, and their dependence on the distance $r_{\alpha\beta}$ between the two ions α and β are fitted from first-principles results [6]. As a further approximation, a minimal basis set is usually adopted: four basis functions (s, p_x, p_y, p_z) per atom are known to be sufficient for a satisfactory description of the valence bands in the case of elemental semiconductors. In order to obtain the single-particle energies ϵ_n and the eigenvectors $c_{l\alpha}^n$ it is necessary to solve the secular problem at each MD time-step. This implies repeated diagonalization of the matrix h, which introduces the $O(N^3)$ scaling of this TBMD formulation. The rank of the matrix is determined by $N = \mathcal{N} * n_b$ where $\mathcal{N}$ is the number of atoms in the simulated systems and n_b is the dimension of the basis set ($n_b = 4$ in the case of covalent semiconductors). Because of current hardware and software limitations, the maximum rank of the matrix that can be

diagonalized on a workstation, provided with a large RAM, is of about $N = 2000$, thus implying a maximum system size of about 500 atoms.

Once the eigenvalues and eigenvectors are known the attrattive potential energy can be computed and summed to the repulsive part derived from a many-body approach [7].

The main difficulties in the parallelization of such codes are: (1) the system cannot be mapped on a regular grid (atoms can move inside the simulation box); (2) in the case of the $O(N^3)$ code a diagonalization of a large sparse matrix is involved (parallel math libraries for sparse matrix diagonalization are not yet available).

$O(N)$ TBMD code

The $O(N)$ scheme which will refer to is that developed by Colombo and Goedeker [4]. The main point of this approach consists in expressing the band-structure energy as $E_{bs} = \text{Tr}[HF]$ where F is the Fermi matrix $F_{\mu,T}$ $F_{\mu,T} = f[(H - \mu)/kT]$ and f is Fermi-Dirac distribution. Given that the trace operation scales linearly, the main burden of this TBMD implementation lies on the calculation of the Fermi matrix. The Fermi matrix can be approximated by a sum up to order n_{pl} of Chebyshev polynomials t_j with h as the argument

$$F_{\mu,T} \simeq \frac{c_o}{2} + \sum_{j=1}^{n_{pl}} c_j t_j(h) \tag{3}$$

where c_j are the coefficients of the expansion in Chebyshev polynomials of the erfc function that well approximates the Fermi-Dirac distribution shape.

In order to evaluate each column of the Fermi matrix, n_{pl} matrix-vector products between h and $|\phi_{l\alpha} >$ are required. It is important to note that the columns of the Fermi matrix are fully independent on each other and can thus be computed in parallel.

At this stage the TBMD code has scaling $O(N^2)$ because the computational complexity is proportional to $(n_b \mathcal{N})^2 n_{pl} n_{off}$, where n_{off} is the number of non-zero off-diagonal elements of the Fermi matrix.

The linear scaling is a consequence of the observation that, in strongly covalent materials, the atomic orbitals are spatially localized with fast exponential decay as a function of the interatomic distance $r_{\alpha\beta}$. Then, the matrix is also sparse with only m_{loc} non-zero off-diagonal elements, where m_{loc} is the number of atoms inside the localization region. When this physical cut-off is introduced, the computational complexity to calculate the Fermi matrix becomes proportional to $n_b m_{loc} n_{pl} n_{off} \mathcal{N}$, hence the $O(N)$ scaling of this TBMD formulation.

Once the Fermi matrix has been obtained, the band-structure energy is calculated as

$$E_{bs} = \text{Tr}[HF] = \sum_{l\alpha} < H\phi_{l\alpha}|F|\phi_{l\alpha} > = \sum_{l\alpha} < H\phi_{l\alpha}|f_{l\alpha} > \tag{4}$$

This calculation can be performed independently for each atom α. Hellman-Feynman forces can be easily obtained from this expression. The repulsive potential contribution is calculated in the same way as in the $O(N^3)$ TBMD formulation.

PARALLEL COMPUTERS

In the following we will review the most important classes of high performance parallel computers [8], namely MIMD (Multiple Instructions Multiple Data), SIMD (Single Instructions Multiple Data) and a new hybrid MIMD-SIMD architecture which has been realized in the frame of a collaboration between ENEA [9] and QSW (Quadrics Supercomputers World) [10].

Parallel computers with MIMD architecture

A MIMD computer consists of a network of processors that are loosely synchronized and can work, at the same time, on different chunks of a code. MIMD computers can be further divided into distributed memory (DM) and shared memory (SM) architectures.

In DM-MIMD computers the processors, each having its local RAM, are connected via a fast network or a switchboard. In SM-MIMD computers (also called Symmetric Multi-Processors, SMP), in turn, the processors are grouped into clusters and share a unique, very large bank of RAM, typically several GBytes. DM-MIMD computers require explicit message (MP) passing libraries (such as MPI [11] and TCGMSG [12]) to perform the allocation and exchange of data between different processors; such operations are not necessary on SM-MIMD computers because the memory is in common among the processors.

Differently from DM-MIMD platforms, where the computational load distribution requires the explicit distribution by the programmer of the data and tasks among the processors, in the SM-MIMD platforms, due to the availability of a shared memory, the compiler is capable of automatic parallelization. The strategy adopted by the compiler consists in dividing the tasks in threads, each thread being automatically sent to a different processor, which "concurs" on the same memory board.

The SM architecture can be emulated on a DM computer by using specific software such as the High Performance Fortran (HPF) [13] and the Global Array (GA) library [16].

HPF is essentially a Fortran 90 with parallelization directives oriented to ease data distribution among the different processors. Several versions of HPF are available, either public domain or proprietary. All of them are related to the standard developed by the International Forum for HPF (HPFF) [13].

The Global Array (GA) toolkit, developed at the Pacific Northwest National Laboratory (Richland, WA) provides a machine-independent, portable "shared-memory" programming interface for distributed memory computers. Each process in a MIMD environment can asynchronously access logical blocks of physically distributed arrays without need for explicit cooperation by other processes.

Parallel mathematical libraries are usually available on MIMD platforms. They can be called either in a code using explicit message passing or in codes written with a HPF or GA. Such libraries are available on the market (NAG [14], IBM PESSL [15]) or public domain (ScaLAPACK, PEIGS).

Parallel computers with SIMD architecture

In a SIMD computer the processors are arranged in a tightly synchronized network and must perform, at the same time, the same instruction on different values of the variables. Because the processors of these machines operate in lock-step, i.e., each processor executes the same instruction at the same time (but on different data items), no external synchronization among processors is required. This greatly simplifies the design of such systems. A control processor issues the instructions that must be executed by the floating-point processors in the processor array. All the currently available SIMD machines use a front-end processor to which they are connected by a data-path to the control processor.

A good example of this class of computers is the Quadrics/APE100 [9, 10], which is currently used in a wide variety of computational fields and is produced by Quadrics Supercomputers World (QSW) [10]. The floating-point processors are arranged on a three-dimensional mesh with periodic boundary conditions, and are synchronized by a control processor. Quadrics can be programmed by an easy-to-use proprietary parallel programming language, called TAO: a Fortran-like language with instructions for data communication among the nearest-neighbour processors. This platform stands out for easy programming, good scalability and peak computational power (100 GFlops in the maximum configuration).

The next generation, called APEmille, will be available in the middle of 1998 [10]. The APEmille is based on the Quadrics architecture of which improves some key features (the local integer addressing will be available). The processing nodes are also much improved, being capable of peak powers of 528 MFlops (66 MHz) or 800 MFlops (100 MHz). The latter will be the building block for the maximum configuration, constitued by 32x8x8 nodes resulting in a peak computational power of 1.6 TeraFlops.

Parallel computers with MIMD-SIMD architecture

In view of extending the capabilities of the TBMD approach to very large scale systems, one should be able to fully exploit the benefits of both MIMD and SIMD architectures.

MIMD computers stand out for the flexibility of use. Indeed, they have been used in many different applications with a good computational efficiency and a relatively small programming effort. Both message-passing libraries and high-level programming languages can be used as simple extension of the standard Fortran and C tools. However, the performances that can be obtained on DM-MIMD computers, on one hand, are limited by the typically complex arrangement of interprocessor communications. Today experience tells that the saturation limit, where a cross-over occurs between the communication overhead and the actual computation time, is reached, depending on the particular application, well below machine sizes of the order of 100 processors. Given the typical computing power of RISC processors, in the range of $\sim$500 MFlops, the practical peak performances on MIMD computers cannot be expected to exceed a few tens of GFlops. On the other hand, SM-MIMD computers, which in principle do not require explicit interprocessor communications, are limited by technological constraints to relatively small machine sizes, typically up to 64 processors.

Compared to MIMD computers, the SIMD can easily be scaled up to extremely high computational power. For example, the Quadrics is already capable of peak powers of 100 GFlops based only on a relatively old-fashion design of the individual processors. The next generation APEmille will be capable of breaking the TeraFlop barrier still with relatively low-price processors. However, SIMD computers usually find their limits in a comparatively more rigid programming scheme, imposed by the synchronized operation of the processors.

In order to get the best of both worlds, a prototype MIMD-SIMD platform has been recently assembled in the ENEA Casaccia Computing Center. The MIMD part of this computer consists of a Meiko-CS2 with eight SUN SPARC nodes and a fast *fat-tree* interconnection network. The SIMD part is represented by 8 Quadris/APE100 towers with 128 to 512 processors each, linked to the MIMD part via fast HIPPI channels (20 MBytes/sec) in such a way to connect each MIMD node to a SIMD partition. This hybrid MIMD-SIMD computer should reach a theoretical peak power of about 75 GFlops. The software employed is the same used on the respective computers plus communication libraries which allow to dynamically interconnect the different portions of the machine, by using Fortran-like I/O calls.

This experimental, hybrid architecture is furthermore ready for the future developments of the individual hardware components: in 1998 we expect to replace the CS2 processors with a QSW Meiko New Machine (called QM1), and the Quadrics/APE100 with APEmille processors; a parallel coordination language is also under development (SKIE) that will allow the compiler to decide the optimal distribution of data and tasks on the MIMD and SIMD parts.

PRACTICAL EXPERIENCES OF PORTING OF TBMD CODES

The operation of matrix diagonalization in the $O(N^3)$ TBMD code requires each processor to know all the elements of the matrix at once. For this reason, a purely DM-MIMD computer is of little help in diagonalization, because the matrix elements are scattered

among the local RAM of each processor and a purely message-passing parallelization would have an exceedingly low efficiency. Then, a solution must be sought in either a true SM-MIMD computer or in a DM-MIMD with SM emulation (i.e., using HPF or GA toolkits).

The $O(N)$ TBMD code, in turn, is relatively easy to parallelize on a DM-MIMD computer. However, in order to demonstrate the capability of attaining the limit of very-large-scale TBMD simulations we will rather present preliminary results of the porting of this code on the high-performance experimental MIMD-SIMD hybrid computer.

Exp. 1: $O(N^3)$ TBMD on the IBM SP2 with HPF

Using the HPF parallel language, the programmer can port his own serial code on a parallel MIMD computer by simply introducing suitable directives for the distribution on the different nodes of the data and the tasks which require the largest computational effort. Thus, the routines parallelized are: calculations of distances among the atoms, construction of the TB matrix, diagonalization of the TB matrix, calculation of the attractive part of the force and calculation of the repulsive part of the force.

In particular, the calculation of the eigenvectors and eigenvalues of the TB matrix is performed in parallel using the routine SYEV of the IBM PESSL mathematical library. To optimize the use of the math routine and because the TB matrix is generally greater ($N *$ $N*8$ Bytes) than the processor RAM (128 MBytes per node on SP2 at ENEA Frascati), the TB matrix is distributed among the processors. As a consequence, the eigenvector matrix and the eigenvector array are also distributed. Furthermore all the atomic coordinates, the attractive and repulsive forces are distributed among the processors.

In this way the target of simulations with a number of atoms of order 10^3 is reached: the parallel code can perform simulations in a reasonable amount of wall clock time and the RAM occupancy of the whole code can be managed.

In Fig.1 is clear a good speedup till about 8 nodes. The point at one processor is real only for the 216 and for the 512-atoms case and extrapolated for the 640 case. Some points corresponding to two processors lie below unity because the time to run on two processors is greater than that to run on a single processor. This is due to the large RAM occupancy of the code when the number of atoms increases.

Exp. 2: $O(N^3)$ TBMD on the IBM SP2 with GA

The GA toolkit allows to emuate a shared memory The matrix to be diagonalized and some workspace arrays are stored in *global* arrays, while all the other arrays and variables are *local* (i.e. replicated) to each processor.

The developed code uses a replicated-data paradigm for the calculations relative to the ionic coordinates in real space (predictor and corrector, time averages of observables, etc.). This part makes use of the standard message-passing TCGMSG libraries [12]. On the contrary, the "electronic" part, i.e., the set up of the hamiltonian matrix, diagonalization and calculation of the Hellmann-Feynman forces, is based on a force-decomposition paradigm (each processor takes care of the construction and operates on a patch of the total matrix, not necessarily delimited by a spatially-contiguous group of atoms). This part is performed entirely within the GA environment. The eigensolver is a general-purpose routine for real-symmetric matrix diagonalization from the parallel PeIGS library [16], which is one of the linear algebra packages compatible with GA.

Fig. 2 shows the results obtained with this implementation of the $O(N^3)$ TBMD code on the IBM SP2 at the MCS Division of the Argonne National Laboratory. The data points are relative to the speedup, expressed as the ratio between the computing time for the scalar code and the wall-clock time for the parallel runs at fixed system size.

It can be seen that also in the case of public domain software the speedup is good allowing to study systems as large as 1000 atoms, reaching the top of the efficiency on 16 nodes.

Exp. 3: $O(N^3)$ TBMD on the DEC4100 with NAG parallel libraries

In this section we would like to underline the usability of parallel mathematical libraries on SM computers. In this case only one processor is in charge for running the application: thus all the routines are performed by one single processor. Parallelization comes about only for the diagonalization of the TB matrix that is done in parallel with a call to the NAG parallel libraries. The philosophy of this subroutine is the following: the whole matrix is in the common memory bank and when the processor reaches this point it activates new processes that automatically take care of a portion of the data. These new processes are spread on all the nodes in the SMP cluster taking into account the total load distribution on the computer. In this case for systems with a number of atoms $\mathcal{N}$ ranging from 216 to 1000, the maximum value for the speedup is at 4 processors.

Exp. 4: $O(N)$ TBMD on the Meiko-Quadrics hybrid MIMD-SIMD

As already stated in the previous sections, the most computationally-intensive part of the $O(N)$ TBMD code is the recursive evaluation of the Fermi matrix. This task can be performed independently for each atom, i.e. in parallel. This part of the code, being the most computational intensive, is performed by the hardware part with the higher computational power, i.e. the SIMD part of the hybrid computer.

The manipulation of the datasets, to be prepared and reorganized for allowing an optimal workload distribution for the SIMD part, is perfromed by the MIMD part. Thus the MIMD part takes care of the input, the inizialization of the arrays, the construction of lists where atoms and their, nearest neighbours and atoms inside the localization region, are stored. The MIMD part reorganizes these data to built a suitable dataset to be sent to the SIMD part. In turn, the SIMD part takes care of the complete calculations of the forces on each atom and of the integration of the equations of motion. At this stage the SIMD part sends back the results to the MIMD part for the calculations of physical averages and the output.

The porting has revealed some interesting results. With the available hardware (CS2 + Quadrics/APE100) the following constraints are been detected in developing the parallel code: (1) all the information must be replicated on the nodes as Quadrics processors cannot handle local integers; (2) for the same reason, localization neighbours and nearest neighbours are indistinguishable; (3) the computational power of single node is too small to perform the task on a number of atoms sufficient to minimize the dataset; (4) scaling with N is no long linear, but slightly worst.

On the contrary, this kind of machine has revealed a large computational efficiency in the evaluation of localized orbitals (12 GFlops sustained on a 512-nodes machine). Furthermore, this work has demonstrated the usability of hybrid architectures to deal with TB simulations. The results of this attempt will be used in the portingto port the $O(N)$ TBMD code on next generation of hybrid machine based on the new SIMD platform APEmille.

CONCLUSIONS

In this paper we have reported the results of the porting of two different implementations of TBMD code on several parallel computers. In particular we have described the porting of the $O(N^3)$ TBMD code on the MIMD computer IBM SP2 with two different software tools (HPF and GA) and the porting of the $O(N^3)$ TBMD code on a Shared Memory computer (DEC4100) using parallel libraries. Furthermore we have showed the porting of an $O(N)$ TBMD code on a hybrid MIMD-SIMD computer where flexibility and high computational power are joined for very large scale simulations.

In particular we can outline the following results:

- both with proprietary parallel language (HPF) or public domain parallel environment a good scalability till 16 nodes is easily reached. It is important to underline that the parallel tools utilized are high level parallel tools. In fact, contrarily to message passing libraries, where the programmer must explicitly manage the inter-node communications, HPF and GA help programmers to decide a general strategy for the distribution of the data among the processors in the parallel computer.

- We have tested the functionality of the porting of the $O(N^3)$ TBMD code on a SM computer using only proprietary NAG mathematical libraries. It is resulted a good scalability and also in this case, where only the diagonalization is performed in parallel, the feasibility of 1000-atoms systems.

- In the case of the $O(N)$ TBMD code we have ported the code on a very powerful MIMD-SIMD computer assessing the limitations induced by the lack of local integer addressing, used very often in the $O(N)$ TBMD code, of the SIMD part. Neverthless, we have obtained good performance in the most computational intensive routine: about 12 GFlops sustained on the 512-nodes SIMD part. This is an optimal base for the porting of this code on the next generation of the hybrid computer composed by a QM1 and an APEmille.

We would like to acknowledge the contribution of drs. B.Di Martino (VCPC, Vienna), M.Briscolini and S.Filippone (IBM Italia) and their support in the development of the HPF+PESSL TB code.

REFERENCES

1. M.P.Allen, D.J.Tildeley, Computer simulations of liquids", Clarendon Press, Oxford (1987).

2. R.Car and M.Parrinello, Phys.Rev.Lett. **55**, 2471 (1985).

3. L.Colombo in: Annual Review of Computational Physics IV, edited by D.Stauffer (World Scientific, Singapore, 1996) p.147.

4. L.Colombo, S.Goedecker, Phys.Rev.Lett. 73, 122 (1994).

5. M.Parrinello and A.Rahman, J.Chem.Phys. **72**, 2662 (1982); S. Nosé and M.L.Klein, Mol. Phys., **50** 1055-1076 (1983).

6. L.Goodwin, A.J.Skinner, D.G.Pettifor, Europhys.Lett. 9, 701 (1989).

7. I.Kwon, R.Biswas, C.Z.Wang, K.M.Ho, C.M.Soukoulis, Phys.Rev.B 49, 7242 (1994).

8. B.Codenotti, M.Leoncini, Fondamenti di Calcolo Parallelo, Addison-Wesley (1990)

9. For more informations refer to the web page www.enea.it/hpcn/hpcn01e.html

10. For more informations refer to the web page www.qsw.com

11. For more informations see the web page: www.mcs.anl.gov/mpi/index.html

12. Informations: www.emsl.pnl.gov:2080/docs/nwchem/tcgmsg.html

13. For more informations see the web page: www.crpc.rice.edu/HPFF/home.html

14. For more informations see the web page: www.nag.co.uk

15. For more informations see the web page: www.ibm.com

16. Informations: www.emsl.pnl.gov:2080/docs/global/ga.html

PREDICTION OF STRUCTURE CANDIDATES FOR SIMPLE IONIC COMPOUNDS USING GLOBAL OPTIMISATION

J. C. SCHÖN, M. JANSEN
Institut für Anorganische Chemie, Universität Bonn, Gerhard-Domagk-Str. 1,
D-53121 Bonn, Germany, e-mail: unc419@ibm.rhrz.uni-bonn.de

ABSTRACT

A method is presented that allows, in principle, the prediction of the existence and structure of (metastable) compounds. We show the results of this approach for two examples of binary and ternary ionic compounds that have not been synthesized yet, but should stand a fair chance of being kinetically stable.

INTRODUCTION

The ability to predict the existence of hypothetical compounds under given thermodynamic boundary conditions, and also to develop realistic routes for their synthesis, is a major step towards one of the main goals of chemistry, the detailed planning of the synthesis of new compounds. In this paper, we concentrate on the first aspect of the planning of syntheses, the determination of hypothetical compounds that are capable of existence.

Our approach[1,2] is based on the analysis of the energy landscape of a chemical compound, using global and local optimisation methods, together with algorithms for the investigation of the phase space structure near local minima of the potential energy, and the analysis and characterization of the structure candidates. The current implementation focusses on ionic compounds, using empirical potentials for the evaluation of the energy function in the global optimisation stage, and a Hartree-Fock-algorithm[3] for refinements. We present results for two compounds, Na_3N and Ca_3SiBr_2, that have not been synthesized yet, but should at least be kinetically stable.

STRUCTURE PREDICTION AND ENERGY LANDSCAPES

Energy Landscape Concepts

In principle, the prediction of metastable compounds of a chemical system should consist of two steps: First, we solve the Schrödinger equation of the system, and then we analyze the statistical mechanical properties and the stability of these solutions. Since this direct approach is not feasible in the case of solids, we simplify the problem by separating ionic and electronic degrees of freedom (Born-Oppenheimer-approximation[4]). Thus, we can construct an (electronic ground state) potential energy hypersurface as function of the classical ionic configurations alone. The prediction of metastable compounds consists in the exploration of this energy landscape.

For $T = 0$ K, it is very clear what constitutes a (meta)stable structure: It is the one, and only one, configuration that is associated with a local minimum of the energy hypersurface. In contrast, for $T > 0$ K, each configuration "i" with energy E_i has a non-zero probability

$$p(i) = \exp(-E_i/k_BT)/Z \ , \quad Z = \sum_j \exp(-E_j/k_BT) \tag{1}$$

of being present at a given time. Therefore, a metastable compound corresponds to some (as yet unidentified) region R of the energy landscape, and the structure of "R" is given by the time-average over all the configurations within R. The equilibrium probability that the system will be found in some configuration belonging to R is given by

$$p(R) = \sum_{i \in R} p(i) = \frac{Z(R)}{Z} = \frac{\exp(-F(R)/k_BT)}{\exp(-F/k_BT)} \tag{2.}$$

Here, $Z(R)$ is the partition function for the region R, and $F(R) = -k_BT \ln(Z(R))$ is defined as the "free energy" for structure R. Clearly, finding the most likely structure representing the system at temperature T corresponds to a) identifying appropriate (kinetically stable) regions R and b) calculating the restricted free energy $F(R)$ and finding the region R^* which minimizes $F(R)$. The compound belonging to the region R^* would be termed "thermodynamically stable". However, for $T > 0$ K, it is not at all obvious, how to identify physically reasonable regions R. For this, we need to define kinetic stability of a compound, based on three relevant time scales: the equilibration time $\tau_{eq}(R)$ that is defined as the time the system requires to reach equilibrium inside the region R of the energy landscape, the escape time $\tau_{esc}(R)$, which is defined as the time it takes the system to leave R and move into the rest of the energy landscape, and the time scale on which we perform our experiments (e.g. a powder diffraction measurement), t_{obs}. If now $\tau_{esc}(R) > t_{obs} \geq \tau_{eq}(R)$, then R is associated with a kinetically stable (locally ergodic) compound.

Stepping-Stone Approach to Structure Prediction

Obviously, the determination of such locally ergodic regions R, and the calculation of their local free energy $F(R)$, are very expensive computationally. At low temperatures, a shortcut is possible, if the region R is surrounded by energy barriers of height E_B, since then $\tau_{esc}(R) \propto \exp(+E_B/k_BT)$. This leads to the following recipe for the prediction of metastable compounds: First one finds the local minima of the hypersurface of the potential energy, called "structure candidates". This is followed by the determination of the energy barriers E_B around these minima, in order to judge their kinetic stability. Finally, one can determine the thermodynamically preferred region R^* by computing $F(R)$ for each kinetically stable region R and finding the minimum of $F(R)$.

When implementing this concept, one is faced with the quandary that on the one hand one would prefer to compute the energy very accurately without any free parameters, e.g. using ab-initio methods, while on the other hand the global optimisation procedures needed for the determination of as many as possible of the low-lying minima of the potential energy require many millions of energy calculations. Therefore, we proceed in a modular fashion: First, empirical potentials are used during the global optimisation, where we have been employing the simulated annealing algorithm[5]. Since these parameters might not be completely appropriate for the unknown compound one investigates, one needs to repeat the opti-

misations for slightly changed values of the parameters, in order to check the robustness of the structure candidates[6]. Next, the physically and chemically most reasonable structure candidates are refined using the Hartree-Fock program[3] CRYSTAL92. This is followed by the determination of the energy barriers around the most important local minima on the energy landscape using the threshold algorithm[7,8], where, because of the high computational effort, it is again necessary to stay within the emipirical potential approximation. These calculations also provide information about the local density of states and possible entropic barriers. Finally, the structure candidates are analyzed with respect to their geometrical and topological properties[9], in order to allow a comparison with known structures in possibly related compounds.

Within the ionic model, the only input into these potentials are average[6] or environment dependent[10] ionic radii, and the hardness of the ions. The approximate potentials for the description of ionic systems consist of three terms, a screened Coulomb term, $\exp(-\alpha r)/r$, a repulsive term, (r^{-n}, n = 12, usually), and an attractive dispersion term, (r^{-6}). If the convergence factor in the Coulomb-term is not present ($\alpha = 0$), the energy function is evaluated using the summation method suggested by deLeeuw[11].

$$V_{ij}(r_{ij}) = \frac{q_i q_j \exp(-\alpha r_{ij})}{4\pi\varepsilon_0 r_{ij}} + \varepsilon_{ij}^1 \left(\frac{\sigma_{ij}}{r_{ij}}\right)^n - \varepsilon_{ij}^2 \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \qquad (3).$$

In addition to these two-body terms, the energy function contains "one-body terms" $E_0(i)$ (the ionisation energy or the electron affinity, respectively), a term pv' when allowing volume changes, where p is the pressure and v' the volume per atom, and the chemical potential $\mu(i)$, the latter being relevant, if the number of atoms is allowed to change during the optimisation. Thus, the energy function per atom we use takes the form

$$\frac{E}{N} = \frac{1}{2N} \sum_{i \neq j} V_{ij}(r_{ij}) + \frac{1}{N} \sum_i E_0(i) \; (+ pv') \; (+ \frac{1}{N} \sum_i \mu(i)) \qquad (4).$$

Furthermore, we use simulation cells with periodic boundary conditions, in order to deal with the large number of atoms in a solid. Note that these periodic boundary conditions are not very restrictive, since size, shape and symmetry of the simulation cell can be freely varied during the many global optimisation runs necessary to achieve a sufficient statistic. In addition to the cell parameters, the location of the atoms and their ionisation state, a number of other parameters can be varied during the optimisation, e.g. the number of atoms and the composition of the system within the cell. The initial configurations for the optimisations usually consist of a cell of ca. 3 - 10 times the volume of all the atoms taken together, with the (neutral) atoms ($N_{atom} \approx 4 - 40$) placed at random positions within the cell.

The determination of the energy barriers employs the threshold-algorithm[7,8], where one performs random walks below a sequence of prescribed energy lids, with periodic quenches, in order to detect the existence of neighbor minima. The geometrical and topological properties of the structure candidates are analyzed using the programs FINDPOLYEDER, FINDCELL and RGS[9].

EXAMPLES

This approach has been successfully tested for a large number of binary and ternary ionic compounds. Here, we will give results for just two examples, the not-yet-synthesized compounds Na_3N and Ca_3SiBr_2.

Na_3N

In table (1), we show the crystallographic data for the seven most promising structure candidates for Na_3N.

Table 1: Most promising structure candidates for the hypothetical compound Na_3N. Cell constants are given in Å. Not-yet-observed structure types are given as (...)-M_3N, where M represents one of the alkali metals.

Structure	Li_3N	Li_3P	$I-Na_3N$	$II-Na_3N$	$I-K_3N$	$III-Na_3N$	Cr_3Si
SG	191	194	139	59	63	221	223
	P6/mmm	P6₃/mmc	I4/mmm	Pmmn(z)	Cmcm	Pm$\bar{3}$m	Pm$\bar{3}$n
Cell	a = 4.5	a = 4.3	a = 4.0	a = 4.34	a = 6.8	a = 3.8	a = 4.84
	b = 4.5	b = 4.3	b = 4.0	b = 5.78	b = 6.1	b = 3.8	b = 4.84
	c = 5.1	c = 7.6	c = 6.75	c = 4.34	c = 5.6	c = 3.8	c = 4.84
Positions	N (1a)	N (2c)	N (2a)	N (2a)	N(4c)	N (1a)	N (2a)
	x = 0	x = 1/3	x = 0	x = 3/4	x = 1/2	x = 0	x = 0
	y = 0	y = 2/3	y = 0	y = 3/4	y = 0.26	y = 0	y = 0
	z = 0	z = 1/4	z = 0	z = 0.22	z = 3/4	z = 0	z = 0
	Na (1b)	Na (2b)	Na (2b)	Na(2b)	Na(4c)	Na (3c)	Na (6c)
	x = 0	x = 0	x = 1/2	x = 3/4	x = 1/2	x = 0	x = 1/4
	y = 0	y = 0	y = 1/2	y = 1/4	y = 0.31	y = 1/2	y = 0
	z = 1/2	z = 1/4	z = 0	z = 0.18	z = 1/4	z = 1/2	z = 1/2
	Na (2c)	Na (4f)	Na (4d)	Na(4c)	Na(8e)		
	x = 1/3	x = 1/3	x = 1/2	x = 3/4	x = 0.29		
	y = 2/3	y = 2/3	y = 0	y = 0.5	y = 0		
	z = 0	z = 0.58	z = 1/4	z = 0.72	z = 0		

When calculating the energy using the program CRYSTAL92, the energetically most favorable structure candidates were Li_3N, Li_3P, Cr_3Si, $I-Na_3N$ and $II-Na_3N$. However, all of these modifications of Na_3N exhibited a slightly positive enthalpy of formation with respect to the elements, suggesting that they might be metastable with respect to decomposition into the elements[12].

Ca_3SiBr_2

In table (2), we show the crystallographic data for four of the most promising structure candidates for Ca_3SiBr_2.

When calculating the energy using the program CRYSTAL92, the energetically most favorable structure candidates were No. 1 and 2. Both these modifications of Ca_3SiBr_2 exhibited a negative enthalpy of formation with respect to the binary compounds Ca_2Si and $CaBr_2$, suggesting that they might be thermodynamically stable against decomposition into the binary precursors.

Table 2: Most promising structure candidates for the hypothetical compound Ca_3SiBr_2. Cell constants are given in Å. The structure candidates are numbered.

Structure	1	2 = a-MgCu$_2$O$_3$	3 = a-NaFe$_2$O$_3$	4
SG	139 I4/mmm	71 Immm	164 P$\bar{3}$m	166 R$\bar{3}$m
Cell	a = 4.5, b = 4.5 c = 19.3 α = 90, β = 90 γ = 90	a = 4.5, b = 6.3 c = 13.5 α = 90, β = 90 γ = 90	a = 4.6, b = 4.6 c = 11.0 α = 90, β = 90 γ = 120	a = 11.5, b = 11.5 c = 11.5 α = 22.8, β = 22.8 γ = 22.8
Positions	Ca (2a) x = 0, y = 0 z = 0 Ca (4e) x = 0, y = 0 z = 0.34 Si (2b) x = 0, y = 0 z = 1/2 Br (4e) x = 0, y = 0 z = 0.83	Ca (2a) x = 0, y = 0 z = 0 Ca (4i) x = 0, y = 0 z = 0.35 Si (2d) x = 1/2, y = 0 z = 1/2 Br (4j) x = 1/2, y = 0 z = 0.17	Ca (1b) x = 0, y = 0 z = 1/2 Ca (2d) x = 1/3, y = 2/3 z = 0.15 Si (1a) x = 0, y = 0 z = 0 Br (2d) x = 1/3, y = 2/3 z = 0.66	Ca (1a) x = 0, y = 0 z = 0 Ca (2c) x = 0.88, y = 0.88 z = 0.88 Si (1b) x = 1/2, y = 1/2 z = 1/2 Br (2c) x = 0.72, y = 0.72 z = 0.72
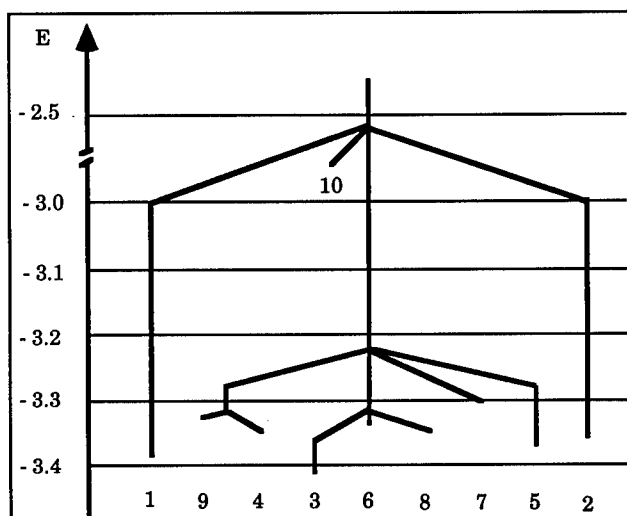
Fig. 1: Tree-graph of the barrier structure of Ca_3SiBr_2. Energies are given in eV/atom. Numbers 1 - 9 refer to single minima, 10 refers to many high-lying minima.

An analysis with the threshold algorithm showed that the structures 1 and 2 were separated by rather high energy barriers from the rest of the energy landscape (Fig. 1), while structures 3 and 4 belonged to a larger basin containing many

local minima that were separted by relatively low energy barriers from each other. The local densities of state for the structures being rather similar, we can conclude that Ca_3SiBr_2 should be capable of existence, and that structures 1 and 2 are the most likely modifications to be found for this compound[10].

CONCLUSIONS

We have presented a general approach to the prediction of crystalline compounds. The current implementation employs an ionic model potential for the construction of the compound's energy landscape that is used for the global optimisation step, and the application of this modular approach to two not-yet-synthesized compounds, Na_3N and Ca_3SiBr_2, has been demonstrated.

Clearly, the choice of the model potential is the most troublesome limitation of the current implementation, since for unknown compounds no a-priori knowledge about their chemical bonding situation is available. Furthermore, in contrast to the typical simulations using empirical potentials[13,14], and related work using optimisation methods in the field of structure determination from powder diffraction data[14,15], the global optimisation requires that the potential must be physically and chemically reasonable everywhere on the energy landscape[1,2,10]. A possible step to ameliorate this situation is the application of computational methods intermediary between the computationally fast empirical potentials and the parameter-free but computationally expensive ab-initio-methods. In particu-lar, tight-binding approaches should be very promising in this respect, because of their general applicability and flexibility. One should note, however, that no symmetry adaption of the wave functions is possible during the explorations of the energy landscape, since the symmetry of the structure candidates is part of the optimisation.

REFERENCES

1. J. C. Schön, M. Jansen, Angew. Chemie (Int. Ed.), **108 (35)**, (1996), 1358 (1286)
2. J. C. Schön, M. Jansen, in Pauling's Legacy - Modern Modelling of the Chemical Bond, Eds. Z. B. Maksic, W.J. Orville-Thomas, (Elsevier, New York, in press)
3. C. Pisani, R. Dovesi, C. Roetti, Hartree-Fock ab-initio Treatment of Crystalline Systems, (Springer, Heidelberg, 1988)
4. J. Callaway, Quantum Theory of the Solid State, (Acad. Press, New York, 1974)
5. S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, Science, **220**, (1983), 671
6. J. C. Schön, M. Jansen, Comp. Mater. Sci., **4**, (1995), 43
7. J. C. Schön, Ber. Bunsenges., **100**, (1996), 1388
8. J. C. Schön, H. Putz, M. Jansen, J. Phys. Cond. Matter, **8**, (1996), 143
9. A. Hannemann, R. Hundt, J. C. Schön, M. Jansen, subm. Acta Crystallogr. A
10. H. Putz, J. C. Schön, M. Jansen, subm. J. Solid State Chem.
11. S. W. deLeeuw, J. W. Perram, E. R. Smith, Proc. Roy. Soc. A, **373**, (1980), 27
12. M. Jansen, J. C. Schön, subm. Z. Anorg. Allgem. Chem.
13. A. M. Stoneham, Handbook of Interatomic Potentials. I. Ionic Crystals (preprint, 1981)
14. C. R. A. Catlow, R. G. Bell, J. D. Gale, J. Mater. Chem., **4**, (1994), 781
15. J. Pannetier et al., Nature, **346**, (1990), 343

COUPLED DYNAMICS OF ELECTRONS AND NUCLEI IN A MOLECULE INTERACTING WITH ULTRASHORT, ULTRA-INTENSE LASER PULSES

S. KHOSRAVI, R.E. ALLEN
Department of Physics, Texas A&M University, College Station, Texas, 77843

ABSTRACT

We have developed a technique for treating the coupled dynamics of electrons and nuclei in a molecule which is subjected to ultrashort and ultra-intense laser pulses. This technique has been employed in quantitatively accurate simulations for H_2^+. Many interesting phenomena have been observed, including photodissociation, bond softening, above-threshold dissociation, ion-population trapping, electron-population trapping, sudden electronic transitions, Rabi flopping, and harmonic generation. Some representative results are shown here.

The interaction of ultrashort and ultra-intense laser pulses with molecules is a current frontier of science. In this paper we introduce a new technique for realistic simulations of the coupled dynamics of electrons and nuclei, when a molecule is subjected to arbitrarily strong electromagnetic radiation. The method involves a quantitatively accurate description of the relevant electronic states and their coupling to the time-dependent electromagnetic potential $A_\mu(t)$. The time-dependent Schrödinger equation (for the electrons) and Newton's equation of motion (for the nuclei) are solved within the interaction picture. Other techniques are used to analyze phenomena such as harmonic generation.

For the initial application to a chemical system, we have chosen H_2^+. This is the simplest molecule in nature, but it exhibits remarkably rich behavior when subjected to an ultrashort and ultra-intense laser pulse [1 − 24]. It is, in fact, an excellent test system for understanding both the standard and the more exotic processes that transpire when a molecule interacts with an intense and time-dependent radiation field.

In our detailed simulations, we have observed many interesting phenomena, including the following [25, 26]:

- **Photodissociation.** The promotion of an electron from a bonding state to an anti-bonding state can produce a repulsive interaction that leads to dissociation.

- **Bond softening.** A very intense laser field deforms each adiabatic potential curve in the vicinity of a multiphoton crossing, causing the lower curve to be flattened or "softened". This effect can produce a dramatic lowering of the barrier to dissociation.

- **Above-threshold dissociation.** The nonlinear effects associated with intense radiation can permit dissociation via various paths: for example, absorption of a single photon, absorption of three photons, or absorption of three photons followed by emission of one photon. The products will thus emerge with a set of different kinetic energies.

- **Ion-population trapping.** The radiation field can create an adiabatic potential well which is deep enough to trap the ions.

- **Electron-population trapping.** An intense laser field can hold an electron in an excited state.

- **Sudden electronic transitions.** These are observed at multiphoton avoided crossings. Even though two dressed-state levels do not cross, there is an exchange of character when they come near each other. As the electron adiabatically follows one curve, it is observed to suddenly undergo either an upward or a downward transition, corresponding to the absorption or emission of one or more photons.

- **Rabi flopping.** In an intense field, an electron can rapidly oscillate back and forth between two accessible states.

- **Harmonic generation.** Photons are radiated at frequencies which are multiples of the fundamental laser frequency.

Because of space limitations, we cannot show all of the above effects in the present paper. Further details of both the method and the results will be presented elsewhere [25, 26]. The main idea is to solve the time-dependent Schrödinger equation (for the electrons in the molecule), together with Newton's equation of motion (for the atomic nuclei), using the interaction picture. Suppose for simplicity that there is only one electron (as there is in the case considered below). One then has the following equations for the coupled dynamics of this electron and the nuclei:

$$
M\frac{d^2}{dt^2}X = -\sum_n \mid c_n(t)\mid^2 \frac{\partial \varepsilon_n}{\partial X} - \sum_{mn} c_m^*(t) c_n(t) \exp(i\varepsilon_{mn}t) \frac{\partial}{\partial X}\langle m \mid V_{int}(t) \mid n\rangle
$$
$$
-\frac{\partial V_{ii}}{\partial X} - \frac{\partial V_{ext}(t)}{\partial X} \tag{1}
$$

$$
i\frac{d}{dt}c_m(t) = \sum_m \langle m \mid V_{int}(t) \mid n\rangle \exp(i\varepsilon_{mn}t) c_n(t) \tag{2}
$$

where $\varepsilon_{mn} = \varepsilon_m - \varepsilon_n$ and

$$
H_{el} \mid n\rangle = \varepsilon_n \mid n\rangle \quad , \quad \mid \Psi\rangle = \sum_n a_n(t) \mid n\rangle \quad , \quad a_n(t) = c_n(t)\exp(-i\varepsilon_n t). \tag{3}
$$

Here X is any nuclear coordinate. The effective classical Hamiltonian for the molecule has the form

$$
H = \sum_j \langle \Psi(t) \mid [H_{el} + V_{int}(t)] \mid \Psi(t)\rangle + V_{ii} + \sum_\ell P_\ell^2/2M_\ell + V_{ext}(t) \tag{4}
$$

where H_{el} is the one-electron Hamiltonian with no field, V_{int} is the electron-field interaction, V_{ii} is the ion-ion repulsion, V_{ext} is the nuclei-field interaction, and the nuclei are labeled by ℓ.

For the present problem of an H_2^+ molecule, detailed considerations [2, 6, 25, 26] give the following results for the energies of the lowest electronic states, and for their interaction with a field of intensity I and frequency ω:

$$\varepsilon_1(R) = 0.1025\{\exp[-1.44(R-2)] - 2.00\exp[-0.72(R-2)]\} \tag{5}$$

$$\varepsilon_2(R) = 0.1025\{\exp[-1.44(R-2)] + 2.22\exp[-0.72(R-2)]\}, \tag{6}$$

$$\langle 1 \mid V_{int}(t) \mid 2\rangle = \gamma\cos(\omega t) \tag{7}$$

$$\gamma = 5.3416 \times 10^{-9}\sqrt{I(W/cm^2)}\mu \tag{8}$$

$$\mu = \langle 1 \mid z \mid 2\rangle = -8.93 + 10\exp[0.0396(R-2)]. \tag{9}$$

All these quantities are given in atomic units, with 0.529 $\mathring{A}$ as the unit of length and 1 *Hartree* as the unit of energy. Under the usual experimental conditions, only these two states (the bonding and antibonding states derived from the $1s$ atomic orbitals) are important, and the only important nuclear degree of freedom is the internuclear separation R.
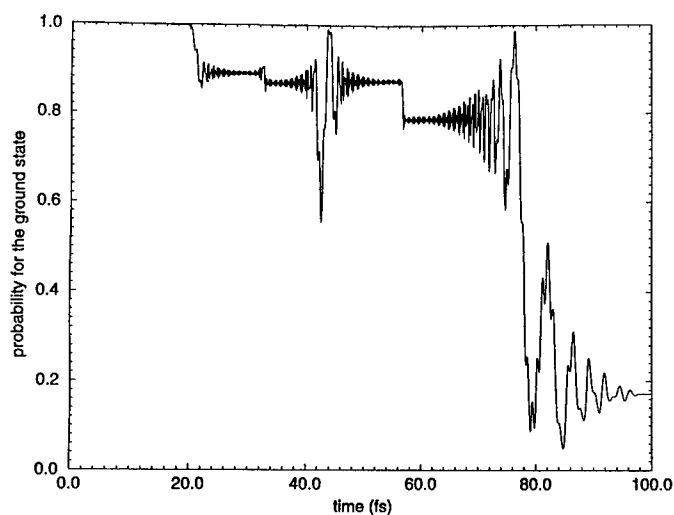
Fig. 1. Square of the amplitude for the ground state as a function of time.

Figs. 1-3 show the results for a simulation in which there was one-photon absorption leading to dissociation. The initial internuclear separation was 1.3 $\mathring{A}$, the initial total energy was -1.95 eV, the photon energy is 2.33 eV, the peak intensity was 1.0×10^{13} W/cm^2, and the pulse duration was 100 femtoseconds. At about 20 fs, the electron becomes partially excited, as can be seen in Fig. 1. There are then rather erratic oscillations in the amplitude $a_1(t)$, or probability $\mid a_1(t)\mid^2 = 1 - \mid a_2(t)\mid^2$, until it becomes dominantly in the excited state $\mid 2\rangle$ at about 80 fs. Since this is an antibonding state, the molecule then dissociates, after executing three vibrations in Fig. 2. The difference between final and initial energies is close to the photon energy of 2.33 eV, as one can see in Fig. 3.

Notice that there is a richness of detail that is missing in conventional treatments based on approximations like Fermi's golden rule. It is also interesting that the usual result of

one-photon absorption emerges automatically, even though the radiation field is treated semiclassically and without the usual approximations. In other simulations, we have observed multiphoton dissociation.
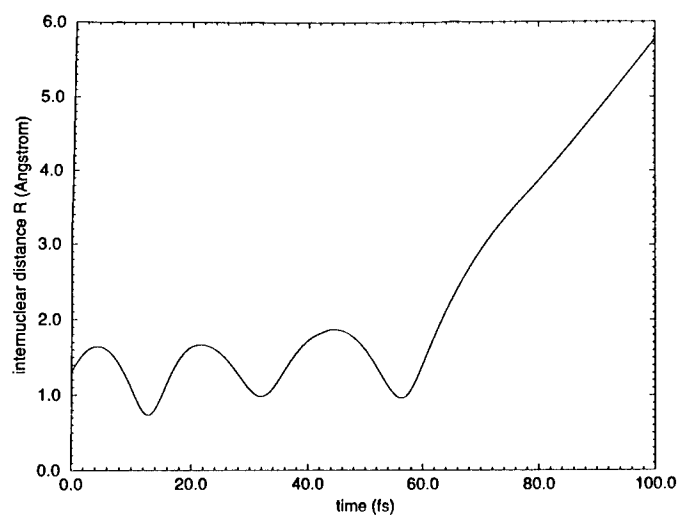
Fig. 2. Internuclear separation as a function of time, demonstrating one-photon absorption.
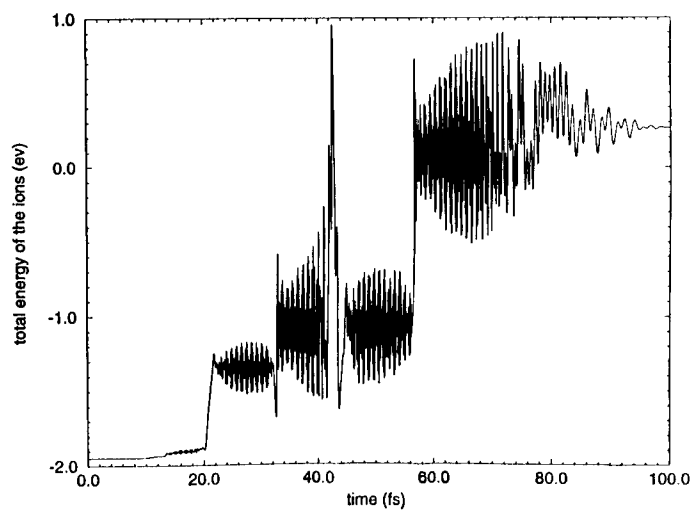
Fig. 3. Total energy as a function of time.

Very unusual behavior has been seen in many simulations. For example, Fig. 4 provides a particularly dramatic example of ion-population trapping: The nuclei exhibit an extremely large excursion from their equilibrium distance, but do not dissociate. Instead, they are trapped by the effective potential energy barrier associated with the adiabatic curves for the dressed states. In this simulation, the molecule was assumed to be prepared in an initial state with an internuclear distance of 2.3 $\mathring{A}$ and an energy of -0.1 eV. The intensity and photon energy were again 1.0×10^{13} W/cm^2 and 2.33 eV, and the pulse duration was 80 fs.
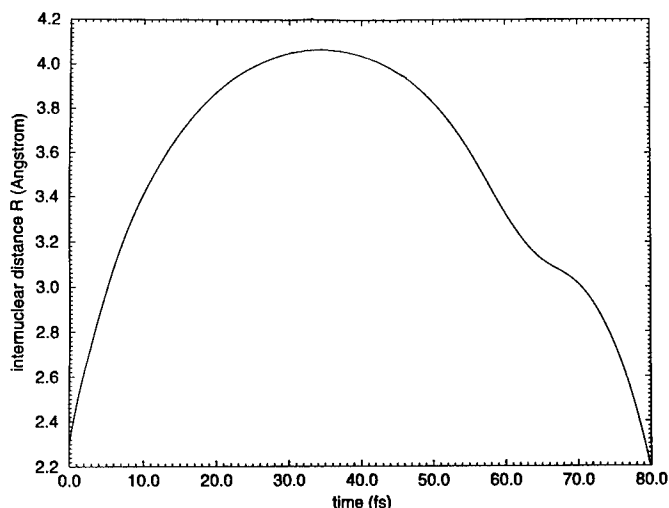
Fig. 4. Internuclear separation as a function of time, demonstrating ion-population trapping in an intense field.

In summary, we have devised a new technique for realistic simulations of the coupled dynamics of electrons and nuclei in an arbitrarily intense radiation field. We have observed many interesting phenomena for the H_2^+ molecule interacting with ultrashort and ultra-intense laser pulses. Two of these are shown here: one-photon dissociation and ion-population trapping.

ACKNOWLEDGEMENT

REFERENCES

1. T. T. Nguyen-Dang, F. Chateauneuf, O. Atabek, and X. He, Phys. Rev. A **51**,1387(1995).

2. X. He. O. Atabek, and A. Giusti-Suzor, Phys. Rev. A **42**, 1585 (1990).

3. Z. Mulyukov, M. Pont, and R. Shakeshaft, Phys. Rev. A **54**, 4299 (1996).

4. M. Chrysos, O. Atabek, and R. Lefebvre, Phys. Rev. A **48**, 3845 (1993).

5. S. Miret-Artes, O. Atabek, and A. D.Bandrauk, Phys. Rev. A **45**, 8056 (1992).

6. X. He, O. Atabek, and A. Giusti-Suzor, Phys. Rev. A **38**, 5586 (1988).

7. A. Giusti-Suzor, X. He, O. Atabek, and F. H. Mies, Phys. Rev. Lett. **64**, 515 (1990).

8. R. W. Heather and F. H. Mies, Phys. Rev. A **44**, 7560 (1991).

9. T. T. Nguyen-Dang and S. Manoli, Phys. Rev. A **44**, 5841 (1991).

10. G. Jolicard and O. Atabek, Phys. Rev. A **46**, 5845 (1992).

11. S. Chelkowski, A. Conjusteau, T. Zuo, and A. D. Bandrauk, Phys. Rev. A **54**, 3235 (1996).

12. A. Giusti-Suzor and F. H. Mies, Phys. Rev. Lett. **68**, 3869 (1992).

13. A. Zavriyev, P. H. Bucksbaum, J. Squier, and F. Saline, Phys. Rev. Lett. **70**, 1077 (1993).

14. P. H. Bucksbaum, A. Zavriyev, H. G. Muller, and D. W. Schumacher, Phys. Rev. Lett. **64**, 1883 (1990).

15. E. E. Aubanel, Jean-Marc Gauthier, and A. D. Bandrauk, Phys. Rev. A **48**, 2145 (1993).

16. O. Atabek and G. Jolicard, Phys. Rev. A **49**, 1186 (1994).

17. T. Zuo, S. Chelkowski, and A. D. Bandrauk, Phys. Rev. A **49**, 3943 (1994).

18. J. Shertzer, A. Chandler, and M. Gavrila, Phys. Rev. Lett. **73**, 2039 (1994).

19. T. Zuo and A. D.Bandrauk, Phys. Rev. A **51**, R26 (1995).

20. N. Moiseyev, M. Chrysos, O. Atabek, and R. Lefebvre, J. Phys. B **28**, 2007 (1995).

21. M. Plummer and J. F. McCann, J. Phys. B **28**, L119 (1995).

22. P. Moreno, L. Plaja, and L. Roso, Phys. Rev. A **55**, R1593 (1997).

23. O. Atabek, M. Chrysos, and R. Lefebvre, Phys. Rev. A **49**, R8 (1994)

24. A. Giusti-Suzor, F. H. Mies, L. F. DiMauro, E. Charron, and B. Yang, J. Phys. B **28**, 309 (1995).

25. S. Khosravi, Ph. D. dissertation (Texas A&M University, August, 1997).

26. S. Khosravi and R. E. Allen, to be published.

ELASTICITY, THERMAL PROPERTIES, AND MOLECULAR DYNAMICS USING NON-EMPIRICAL TIGHT-BINDING

RONALD E. COHEN [*], LARS STIXRUDE [**], AND EVGENY WASSERMAN[***]
[*]Carnegie Institution of Washington, 5251 Broad Branch Rd., N.W., Washington, DC 20015
[**]University of Michigan, Ann Arbor, MI
[***]Battelle, Pacific Northwest National Laboratory, Richland, WA

ABSTRACT

We have further developed and applied a new non-empirical tight-binding total energy model to properties of Si, Xe, and Fe at high pressures. We have studied elasticity of various phases of each of these, demonstrating that the new model is applicable to a wide range of materials, including semiconductors, rare gases, and transition metals. We have used the particle-in-a-cell method to study the thermal equation of state of hcp Fe and find excellent agreement with the shock equation of state. A molecular dynamics code has been developed based on this method, and we have studied the properties of Fe liquid at high pressures.

INTRODUCTION

Self-consistent methods are being applied to ever-increasingly complex problems, including molecular dynamics simulations, but system sizes and number of configurations that can be explored in a reasonable amount of time are still limited. Potential models generally lack what is needed to treat covalent and metallic systems well, except for limited parts of configuration space. Tight-binding total-energy models fit in the niche between potential models and self-consistent methods, being much faster than self-consistent methods and containing the essential physics to treat metallic and covalent systems. We have developed a tight-binding total-energy model which is fast and accurate [1], which we have demonstrated works well for semiconductors (Si), rare gases (Xe), and transition metals (Fe) [2]. The method works well through coordination changes and metallization in Si and Xe, yet there are no explicit structure dependent terms. There is no fitting to experiment; it is first-principles in the sense that all parameters are obtained by fitting accurate Linearized Augmented Plane Wave (LAPW) computations [3]. Here we review the method and some recent results.

METHOD

The method is described in detail in Ref. [2]. In brief the total energy is represented as the band structure energy, with no explicit pair potential terms. This is done by shifting the eigenvalues of the band structure computations so that their sum is the total energy, plus a constant. This procedure removes ambiguity with respect to the energy zero for the band structures, which is structure and volume dependent for extended systems (i.e. periodic boundary conditions). Semi-core overlap can be accounted for with a pair potential using Gordon-Kim[4] for the overlap energy if necessary; we find this necessary at extreme pressures (>300 GPa) in Fe, but may be important in other systems at lower pressures if there is significant semi-core state overlap for states not included explicitly in the tight-binding fit.

We fit the individual off-diagonal interactions to a two-center form

$$P = (a + br)\exp[-c^2 r]\zeta(r) \tag{1}$$

where r is the distance between the two atoms and $\zeta(r)$ is a cut-off function that vanishes beyond some distance. Onsite terms were fit using two different models. In the first case we used

$$D = e + g\rho^{2/3} + h\rho^{4/3} + i\rho^2 \tag{2}$$

where

$$\rho = \sum \exp[-d^2 r]\zeta(r) \tag{3}$$

is the local density around an atom. This model did not work well for Si. In the second case, we accounted for the integrals <i|j|i> for diagonal interactions for a orbital on one atom with a potential on a second atom [5]. The formulation is given in ref. [2]. This second method accounts implicitly for structure dependent shifts in the on-site terms, including e_g - t_{2g} splitting, etc. We simultaneously fit total energies and band structure eigenvalues for a set of structures and volumes. Much better results were obtained using a non-orthogonal basis, though an acceptable orthogonal model for Fe was obtained for more rapid molecular dynamics simulations. The parameters we obtain are effective parameters due to the compromises necessary to obtain accurate energies and band structures with a minimal basis two-center representation over a wide compression range. Thus they may not be in perfect agreement with parameters obtained using an exact inversion scheme [6]. The non-orthogonal parameters for Fe, Si, and Xe are available [7].

RESULTS

Si

We fit 33 different structures and volumes for silicon, including fcc, bcc, hcp, sc, and diamond, using s, p, and d states, with a total of 72 parameters fit to 32841 input data (33 total energies and 32808 eigenvalues) from our LAPW computations. Figures 1 and 2 show the diamond and fcc band structures. The fit for the occupied states is quite acceptable.
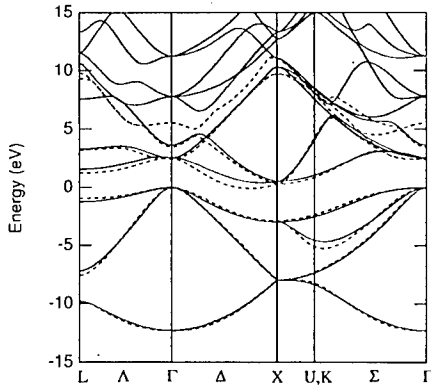
Figure 1. Cubic diamond and structure. Solid lines are LAPW and dashed are the tight-binding fit.
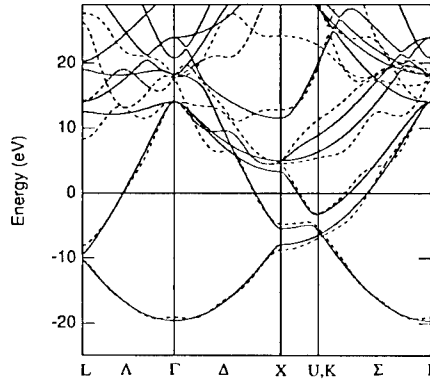
Figure 2. Band structure for fcc Si.
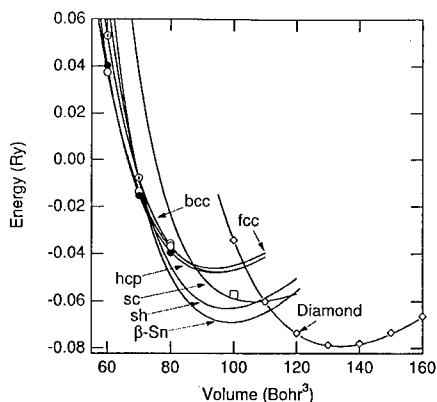
Figure 3. Lines are from the tight-binding model. Points are from LAPW computations.

Figure 3 shows the total energy for different Si structures as functions of volume. The equation of state is reproduced over 62% compression. Elastic constants and the diamond Raman frequency are also predicted well by the model [2].

Xe

At zero pressure xenon is a so-called van der Waals solid consisting of Xe atoms with correlation effects in the overlap giving the binding. Dispersion of the eigenstates is very small. With increasing pressure the bands widen, until 90 GPa in LDA where metallization occurs. We have found that our tight binding parametrization gives this behavior accurately (Fig 4). The equation of state is in excellent agreement with experiment over a four-fold compression (Fig. 5).

Although we used only hcp, fcc, and bcc structures for the fit, we find that the elastic constants agree well with experiment [8], and their pressure dependence is excellent as well (Fig. 6). This is true even though the elastic constants change drastically with pressure in Xe due to its high compressibility.
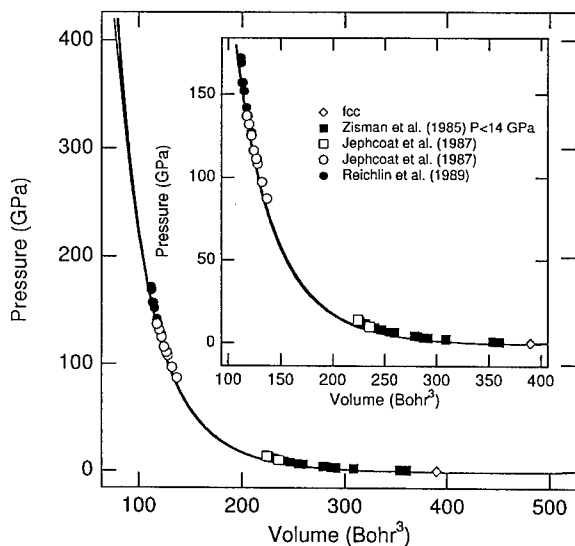
Figure 5. Equation of state for Xe. Lines tight-binding (and LAPW); points, experiment.

Figure 4. LDA band structures for hcp Xe at (a) 0 GPa, (b) 20 GPa, and (c) 250 GPa. Solid lines are LAPW, dashed are tight-binding.

503
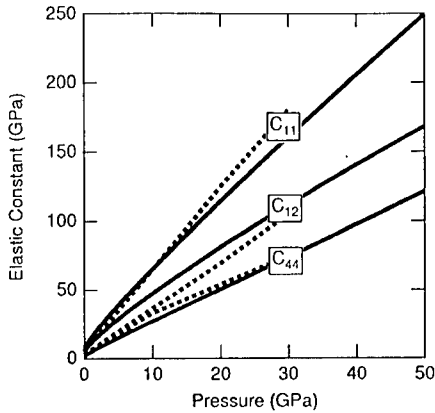
Figure 6. Elastic constants of Xe. Lines are the tight-binding model, curves are from experiment
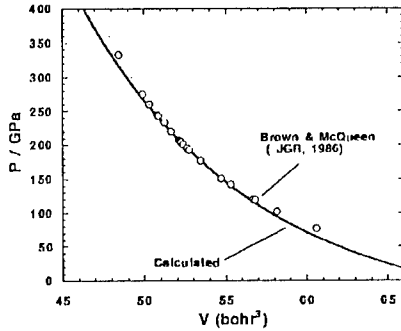
Figure 7. Computed Hugoniot for hcp Fe compared with experiment

relations between atomic motions, and should work very well at high temperatures. We used about 100 atom supercells for most computations, with tests for up to 256 atoms. The electronic contribution to the free energy was obtained from the LAPW band structure, and thermal vibrational contributions from the particle-in-a-cell method. The computed Hugoniot is in excellent agreement with experiment (Fig. 7) [14].

We also have obtained an orthogonal fit for Fe in order to perform

Fe

As the third example we discuss Fe, a transition metal. We used the GGA [9] since LDA does not give the correct ground state and the GGA gives an excellent equation of state [10]. We have not yet developed a spin-polarized version of our parametrization, so we concentrate on the high pressure hcp phase which is not magnetic. A magnetic model could be generated along the lines of Zhong et al. if desired [11], or explicit spin dependence could be included in a significantly more complex model. The band structures we obtain are again in excellent agreement with the LAPW [2].

We have compared the elastic constants obtained using LAPW frozen phonon computations with the tight-binding model, and find excellent agreement, even for strains not directly explored in the input set of structures [12].

Using the tight-binding model for iron, we have computed the high pressure thermal equation of state using the particle-in-a-cell model [13]. The partition function is estimated by integrating over motions of a single atom in a supercell. This includes anharmonicity neglecting cor-
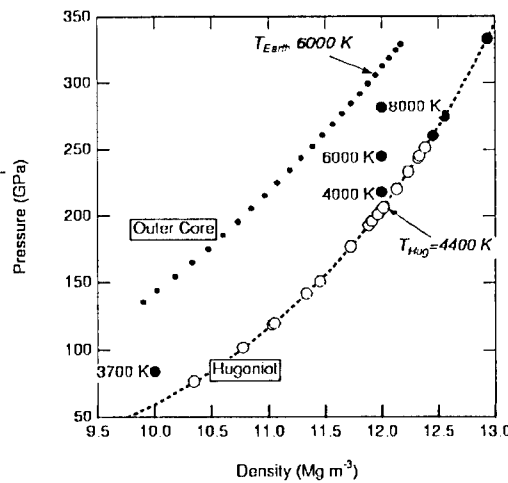
Figure 8. Equation of state points for liquid Fe from MD compared with the hcp Hugoniot and geophysical inversion of the Earth's outer core.

faster molecular dynamics simulations [15]. Figure 8 shows some equation of state points for liquid iron from the MD simulations compared with the Hugoniot for solid hcp iron, and with the Earth's outer core, which is mostly liquid iron but contains some light elements as well.

We have also computed the radial distribution function for liquid Fe under core conditions (Fig. 9). Using this we have developed a model for the viscosity of liquid iron at high pressures. We obtain about 4 centipoise at 4000 K for densities of 10-12 Mg/m^3.
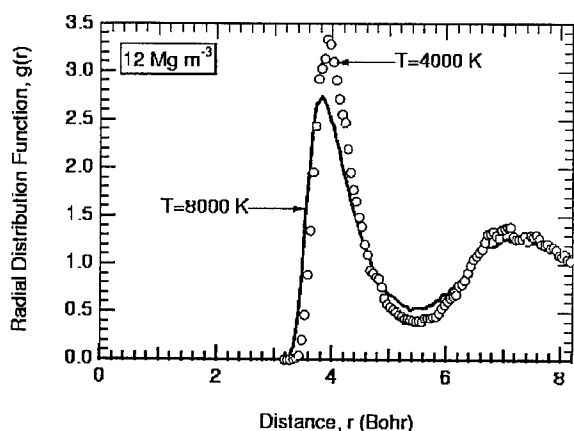
Figure 9. Structure factor for liquid Fe under Earth's outer core conditions.

SUMMARY

We have developed and tested a tight-bonding model that works well for covalent, metallic, and rare-gas systems over wide compression ranges. This model holds promise for fast and accurate simulation of the dynamical properties of large systems.

REFERENCES

1. R.E. Cohen, M.J. Mehl, and D.A. Papaconstantopoulos, Phys. Rev. B, **50**, p. 14694 (1994).

2. R.E. Cohen, L. Stixrude, and E. Wasserman, Phys. Rev. B, **56**, p. 8575 (1997).

3. D.J. Singh, *Planewaves, Pseudopotentials, and the LAPW Method*, pp. 115, Kluwer Academic Publishers, Boston, 1994.

4. R.G. Gordon and Y.S. Kim, J. Chem. Phys., **56** , p. 3122 (1972).

5. J.L. Mercer, Jr. and M.Y. Chou, Phys. Rev. B, **49**, p. 8506 (1994).

7. R.E. Cohen, L. Stixrude, and E. Wasserman, http://granite.ciw.edu/~cohen/research/elastic_tb/sktab.htm .

8. A. Polian, In: *Frontiers of High-Pressure Research*, Eds: H. D. Hochheimer and R.D. Etters, pp. 181-192, Plenum Press, New York, 1991.

9. J.P. Perdew et al., Phys. Rev. B, **46**, p. 6671 (1992).

10. L. Stixrude, R.E. Cohen, and D. Singh, Phys. Rev. B, **50**, p. 6442 (1994).

11. W. Zhong, G. Overney, and D. Tomanek, Phys. Rev. B, **47**, p. 95 (1993).

12. G. Steinle-Neumann, L. Stixrude, and R.E. Cohen, [abs] EOS Trans.Am.Geophys.Union, **in press**, (1997).

13. E. Wasserman, L. Stixrude, and R.E. Cohen, Phys. Rev. B, **53**, p. 8296 (1996).

14. J.M. Brown and R.G. McQueen, J. Geophys. Res., J. Geophys. Res. B, **91**, p. 7485 (1986).

15. L. Stixrude, E. Wasserman, and R.E. Cohen, [abs] Bulletin American Physical Society, **42**, p. 90 (1997).

THE EFFECTS OF THE ELECTRON-PHONON INTERACTION ON THE VIBRATIONAL ANOMALIES AND POLYMORPHISM IN TITANIUM

†‡ J.L. Gavartin and † D. J. Bacon,
†Department of Materials Science and Engineering, The University of Liverpool, Liverpool
L69 3BX, U.K., j.gavartin@liverpool.ac.uk
‡Institute of Chemical Physics, University of Latvia, LV1586 Riga, Latvia

Abstract

We apply the frozen phonon and molecular dynamics methods within the semiempirical orthogonal tight-binding framework to study the anomalous behaviour of the (0001) optical longitudinal (LO) and transverse (TO) phonons in the low temperature hcp phase of Ti, and the $\frac{2}{3}[111]L$ and $\frac{1}{2}[110]T_1$ phonons in the high temperature bcc phase. We demonstrate that, in agreement with previous findings in Zr, the anomalous thermal frequency shifts in hcp Ti are related to the strong coupling of the electron density of states (DOS) to the particular lattice distortions. The distortions along the bcc $\frac{2}{3}[111]L$ and $\frac{1}{2}[110]T_1$ phonons also significantly affect the DOS, resulting in the instability of these modes at low temperatures and triggering the bcc-hcp and bcc-ω phase transformations.

Introduction

The group IV transition metals (titanium, zirconium and hafnium) possess a number of unusual properties attributed to strong anomalies in their phonon spectra. These include crystalline polymorphism, a high temperature saturation of the electrical conductivity, rapid increase of the constant pressure heat capacity at high temperature, and marked anisotropy in diffusion and thermal properties. The vibrational anomalies manifest themselves by a strong (and sometimes inverse) temperature dependence of certain phonon frequencies, even at conditions far away from the phase transitions. For example, almost all the $[0001]LO$ branch in the low temperature hexagonal α-phase of Ti and Zr exhibits substantial positive frequency shift with increase of temperature [1], while the frequencies of the TO branch decrease to an extent much larger than predicted by thermal lattice expansion. In the high temperature bcc-structured β-phase the $[110]T_1$ zone-boundary and $[111]L$ ($q = \frac{2}{3}$) phonons are found to have frequencies below 1THz in the vicinity of the $\beta - \alpha$ transformation (1055 K in Ti) and to be strongly damped [2]. Further, the frequency of the $[110]T_1$ zone boundary phonon is found to increase with temperature over the whole region of existence of the β-phase (up to 1930 K in Ti), while the frequency of $\frac{2}{3}[111]L$, although overdamped, does not show appreciable temperature dependence [2].

Varma and Weber [3] suggested that the occurrence of the phonon anomalies in transition metals is connected to the sensitivity of the electronic bands near the Fermi-level to the particular lattice distortions, causing a reduction in electronic energy. This idea was further developed by Liu and co-authors [4] to the $[0001]LO$ phonon anomaly in α-Zr. Using the frozen phonon (FP) approach within the local density functional theory they demonstrated that the lattice distortion along this phonon lifts the degeneracy of the energy bands in the Brillouin-zone boundary normal to the (0001) direction, thus resulting in a decrease of the density of states at the Fermi-level and LO phonon anharmonicity.

The FP calculations for β-Zr [5] revealed the intrinsic instability of the bcc phase relative to the $\frac{1}{2}[110]T_1$ and $\frac{2}{3}[111]L$ phonons, which are correspondingly linked to the β to α and β to the high pressure ω phase transformation mechanisms.

In the present paper we study the phonon anomalies in α and β phases of titanium using

the orthogonal tight-binding approach. We demonstrate the consistency of the phonon frequencies obtained by molecular dynamics and FP calculations. We also analyse how the phonon properties are coupled to the details of the electronic spectrum of the system.

Calculation procedure

We employ the frozen phonon method [4,5] for the calculation of phonon frequencies and third- and forth-order phonon force constants, and in the case of α-Ti we compare these results with the frequencies obtained from (NVE) molecular dynamics simulations by evaluating the power spectrum of the projected velocity autocorrelation function [6]. Both methods assume Born-Oppenheimer (full adiabatic) approximation, in which the electronic density of states instantly follows the changes in atomic positions. For the energy and force calculation we use the semiempirical orthogonal tight-binding (TB) method within the k-space formalism and within the O(N) bond order potential (BOP) method implemented in the OXON code [7]. The parameters for the hopping integrals and the pairwise repulsive part were developed by Girschik and co-authors [8], and have been fitted to reproduce the equilibrium lattice and elastic constants for α-Ti. Originally TB parameters were fitted with fixed number of moments within BOP formalism (9 moments in the electron DOS expansion) and finite electron temperature of 0.3 eV [8]. In order to control the effects of finite number of moments in the DOS evaluation, we have tested the parametrization within the k-space method and zero electron temperature, and found it to give similar cohesive energies and static geometries to BOP for α-, β-, fcc, and ω-phases of titanium. The static energies give the following sequence in stability of titanium polymorphs: α (ground state), ω (almost degenerate to α), fcc, β, which is in agreement with the full potential muffin-tin local density functional calculations. However, the calculated static lattice constant for β-Ti, $a_\beta = 3.20$ Å is lower than the experimental value $a_\beta^{exp} = 3.28$ Å. This results in the predicted density of the β-phase being the same as of the high-pressure ω-phase, while experimentally the latter is higher. This discrepancy is due to the absence in the TB scheme of the explicit environmentally-dependent terms, and thus a poor account for the variation of interactions with a change in local coordination. Having these limitations, the TB parametrization developed in [8] reproduces reliably the static and, as we show in this paper, important dynamic properties of different titanium polymorphs. In the BOP method we have used 13 moments in the DOS expansion and an electron temperature 0.1 eV. We also show in this paper that this level of approximation has little effect on the accuracy of the evaluated phonon frequencies and higher order force constants.

As the crystalline lattice distorts, so does the charge density, causing the changes in bonding energy. In tight-binding calculations this is reflected by a change of the electronic spectrum or electron density of states with respect to atomic positions. The important characteristics reflecting the changes in the DOS (which also enter any theory of the electron-phonon interaction) are the position of the chemical potential μ (= Fermi-level at T=0 K), and the density of states at the chemical potential level $n(\mu)$.

Phonon calculations

Let us first consider the optical phonons in titanium α-phase. At q=0 the displacement along the longitudinal phonon consists of the antiphase vibration of the adjacent basal planes along the [0001] axis. The transverse optical phonon ($TO_{\parallel}$) can be thought of as the vibration of the basal planes along some axis normal to the [0001] direction. We consider here the TO phonon with [01$\bar{1}$0] polarization. The zone boundary (0001) phonons ($q = \frac{1}{2}$) consist

508

Table 1: The frequencies of the transverse and longitudinal (0001) phonons of α-Ti and the third- and forth-order force constants calculated at the zone centre and zone boundary within the frozen phonon and MD (96 atoms) methods. The experimental data are taken from [1]. The frozen phonon curves are calculated with 13 moments in electron DOS expansion and electron temperature 0.1 eV. In brackets are given the values calculated with 41 moments and kT=0.001 eV.

Phonon		Frequency (THz)				g	h
		Frozen	MD		Exp.	(eV $\mathring{A}^{-3}$)	(eV $\mathring{A}^{-4}$)
		0 K	59 K	292 K	295 K		
LO	(q=0)	5.90 (5.75)	5.90	5.92	5.54	0.0	336.0 (316.8)
TO	(q=0)	4.48	4.25	3.74	4.10	4.2	-307.2
LO=LA	(q=0.5)	3.87 (3.84)	4.0	3.82	5.73	0.0	-31.2 (-67.2)
TO=TA	(q=0.5)	2.57	2.3	2.05	3.05	0.0	-264.0

of the corresponding antiphase vibrations of even basal planes, while the odd planes stay frozen. First we have calculated the adiabatic curves of cohesive energy $U(\mathbf{Q}_k)$ against the displacement vectors $\mathbf{Q}_k$ corresponding to LO and TO phonons at the zone centre and zone boundary up to the maximum amplitude 0.1 $\mathring{A}$, (index k numerates both a phonon branch and a wave vector). Then, in order to obtain the force constants, we approximated each curve $U(\mathbf{Q}_k)$ by its least-square fit polynomial function:

$$\tilde{U}(\mathbf{Q}_k) = U_0 + \frac{1}{2}\omega_{0k}^2 M Q_k^2 + \frac{1}{6}g_k Q_k^3 + \frac{1}{24}h_k Q_k^4. \tag{1}$$

Here U_0 is the energy at equilibrium, ω_{0k}, g_k, and h_k are respectively the harmonic frequency, third-, and forth-order force constants corresponding to a phonon k, M is the atomic mass of titanium. The frequencies ω_0 are listed in the table 1 together with the experimental frequencies obtained in inelastic neutron scattering experiments [1] and from MD simulations. The calculated values of higher order force constants g_k, and h_k are also shown in the table.

Next we consider a longitudinal $\frac{2}{3}[111]$ phonon in β titanium. The extension of this phonon to finite displacements leads to the phase transformation from β to the high pressure ω-phase. This transformation can be seen as a distortion of the ABC... stacking of the (111) bcc planes, when B and C planes are moving towards each other along the [111] axis, whereas A planes stay at rest. The hexagonal ω phase with three atoms in the unit cell is obtained when B and C planes collapse into one plane for a displacement $a_\beta \frac{\sqrt{3}}{12}$. The adiabatic cohesive energy curve along the $\frac{2}{3}[111]L$ phonon displacement is depicted in figure 1 (top). The bottom curve represents a corresponding shift of the chemical potential.

Finally, we consider the zone boundary $\frac{1}{2}[110]T_1$ phonon (N-point phonon), which is a first step in the bcc-hcp transformation along the Burgers path. The finite displacement along this phonon produces the hcp stacking by the antiphase displacement of the two neighbouring ($1\bar{1}0$) planes by the distance $a_\beta \frac{\sqrt{2}}{12}$ in the [110] direction. The hcp structure is then obtained by two equivalent long-wave shears deforming irregular hexagons within the ($1\bar{1}0$) planes into the regular ones, so the ($1\bar{1}0$) planes in the bcc structure become the (0001) basal planes in the hcp structure. The cohesive and bonding energy curves together with a chemical potential shift corresponding to the $\frac{1}{2}[110]T_1$ phonon are shown in figure 2.
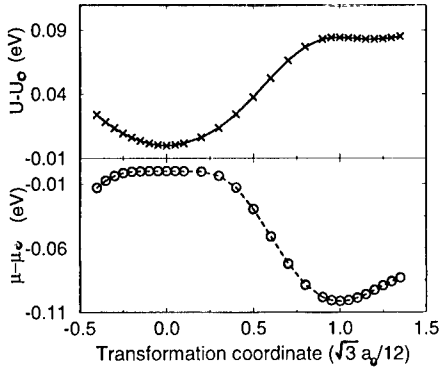
Figure 1: The cohesive energy (top) and chemical potential (bottom) versus displacement along the $\frac{2}{3}[111]L$ phonon. 0 corresponds to the ω phase 1 to the bcc phase; a_0 is the bcc equilibrium lattice constant. U_0 - cohesive energy of bcc, μ_0 - chemical potential of bcc.
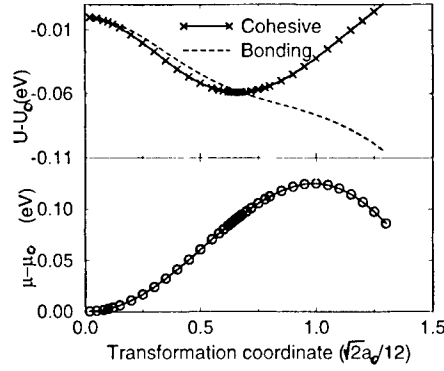
Figure 2: The cohesive and bonding energy per atom (top), and chemical potential (bottom) versus displacement along the $\frac{1}{2}[011]T_1$ phonon. 0 corresponds to the bcc structure; 1 to the hcp stacking of the (110) planes.

Discussion

LO and TO phonons in $\alpha - Ti$

As seen in table 1, the calculated FP and MD phonon frequencies are in good agreement with the experimental data [1], although a substantial difference is observed for the zone-boundary longitudinal phonon. Also, FP frequencies agree well with the values obtained from the MD simulations. The comparison of the MD frequencies calculated at temperatures 59 and 292 K show that, in agreement with the experiment [1], the zone center LO phonon becomes stiffer with temperature, while the TO phonon becomes softer. At the zone boundary the frequencies of both longitudinal and transverse phonons decrease with temperature. The qualitative information on the thermal frequency shifts can be also obtained from the frozen phonon calculations using the following argument. Neglecting the effects of phonon-phonon coupling, each phonon can be considered as a one-dimensional conservative anharmonic oscillator with the potential function (1). Then, if the forth-order force constant h_k is positive, the potential (1) will be steeper than the reference harmonic potential $\frac{1}{2}\omega_{0k}^2 MQ_k^2$, thus resulting in frequency increase with the oscillators energy (temperature). Conversely, $h_k < 0$ (with any sign of g_k) results in the potential (1) being softer then the reference harmonic potential, that is, the corresponding frequency will decrease with temperature. Given these considerations our FP calculations (table 1) predict at the zone centre strong positive thermal frequency shift for the LO phonon, and negative shift for TO phonon. At the zone boundary both (0001) longitudinal and transverse phonons display the negative frequency shifts. We note that the lattice thermal expansion has not been considered in either FP or MD calculations. Since Ti has normal (positive) thermal expansion coefficient, the account of this effect would weaken the positive and enhance the negative frequency shift.

In order to demonstrate the effects of the lattice distortions on the electron distribution,
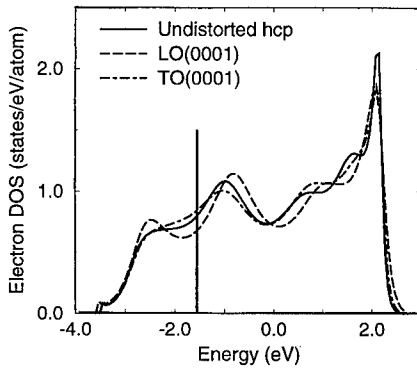
Figure 3: The electron density of states calculated for the equilibrium α-Ti (solid line), and for the lattice distorted along the zone centre LO (dashed line) and TO (dot-dashed line) phonons. The displacement is 0.1 $\mathring{A}$. The vertical lines indicate the positions of the Fermi-level.

we analyse the electron DOS corresponding to a fully relaxed α-structure, and to 0.1 $\mathring{A}$ atomic displacements along the LO and TO zone-centre phonons (figure 3). In agreement with [3,4], the LO phonon indeed causes the marked decrease of the density of states at the Fermi-level $n(\mu)$. The position of the Fermi-level for both phonons shifts slowly to the lower energies with amplitude increase. Another remarkable feature of the distorted density of states for the LO phonon is that it acquires significant changes in the energy region as far as 1 eV below the μ. This suggests that some of the low-lying energy bands are strongly coupled to the LO phonon as well as the earlier suggested [4] bands at the Fermi-level. Therefore, the semiquantitative model proposed in [4], in which the electronic energy change causing the frequency shift is associated solely with the change in the few selected bands in the vicinity of the Fermi-level, is somewhat oversimplified. In contrast to the longitudinal phonon, the distortion along transverse phonon results in the increase in the $n(\mu)$ (figure 3). This causes anomalously strong thermal softening of the TO branch reported for both α-Ti and α-Zr [1].

We note that, although the details of the density of states depend on the number of moments used in the DOS calculation, the same trends have been observed in the DOS approximated by 41 moments or in k-space calculations with electron temperature 0.001 eV.

$\frac{2}{3}[111]L$ and $\frac{1}{2}[110]T_1$ phonons in $\beta - Ti$

Our FP calculations in β-Ti predict the imaginary frequencies for both $\frac{2}{3}[111]L$ and $\frac{1}{2}[110]T_1$ phonons (figures 1,2) indicating the instability of the titanium bcc phase at 0 K towards the bcc-ω and bcc-hcp transformations. The interesting feature of the bcc-ω transformation is that the electron DOS changes substantially with the small lattice distortions around the bcc structure, but it is almost completely insensitive to the distortions near the ω-phase. This is reflected by both the change of the chemical potential μ (figure 1 (bottom)) and $n(\mu)$ (not shown) with the transformation coordinate. This suggests that the ω-Ti may also have a trigonal modification (when B and C planes are not totally collapsed) along with the hexagonal one discussed above. Such a structure has been recently observed in Ti films grown on an iron surface [9].

It is seen in figure 2 (bottom), that the chemical potential curve for the $\frac{1}{2}[110]T_1$ phonon displacement has two extrema - the minimum in the bcc structure and the maximum corresponding to the hcp stacking of the bcc (011) planes. We note that chemical potential reflects the changes in the electronic Gibbs free energy [10]. The more detailed analysis

shows that the hcp stacking configuration corresponds to the minimum of the Gibbs free energy along the chosen direction, which results from the interplay between the bonding energy (see figure 2) and electronic [pressure × volume] term. The pair repulsive part of the potential counterbalances the decrease in bonding energy and causes the shift of the *total* energy minimum somewhere in between bcc and hcp stacking. Therefore, the transformation along the Burgers mechanism, requires a simultaneous (not consequtive) action of the N-point phonon and long-wave shears. Such a process is only statistically important when the frequency of the N-point phonon falls to zero, in order it to 'freeze' and interact with low frequency shear deformations.

In summary we have modelled the phonon anomalies in α- and β-Ti and found the simple orthogonal TB approximation to be adequate in describing the effects of the electron redistribution coupled to the particular lattice distortions. Having their own limitations, the FP and MD methods complement each other in calculations of the lattice dynamics and the scheme as a whole makes it a powerful tool to further studies of defect properties which are currently in progress.

Acknowledgements

References

1. C. Stassis, D. Arch, B.N. Harmon, and N. Wakabayashi, Phys. Rev. B **19**, p. 181 (1979). - Titanium; C. Stassis, J. Zarestky, D. Arch, O.D. McMasters, and B.N. Harmon, Phys. Rev. B **18**, p. 2632 (1978). - Zirconium.
2. W.Petry, A. Heiming, J.Trampenau, M.Alba, C.Herzig, H.R. Schober, and G.Vogl, Phys. Rev. B **43**, p. 10933 (1991) - Titanium; A. Heiming, W.Petry, J.Trampenau, M.Alba, C.Herzig, H.R. Schober, and G. Vogl, ibid., B **43**, p. 10948 (1991) - Zirconium.
3. C.M. Varma and W. Weber, Phys. Rev B **19**, p. 6142 (1979).
4. S.H. Liu, C. Stassis, and K.-M. Ho, Phys. Rev. B **24**, p. 5093 (1981).
5. K.-M. Ho, C.L. Fu, B.N. Harmon, W. Weber, and D.R. Hamann, Phys. Rev. Lett. **49**, p. 673 (1982); K.-M. Ho, C.L. Fu, and B.N. Harmon, Phys. Rev. B **29**, p. 1575 (1984); Y.Y. Ye, K.M. Ho, B.N. Harmon, and P.-A. Lindgard, Phys. Rev. Lett. **58**, p. 1769 (1987).
6. J.L. Gavartin, D.J. Bacon, Comp. Materials Science, (1997) (in print).
7. A.P. Horsfield, A.M. Bratkovsky, M. Fearn, D.G. Pettifor, M. Aoki, Phys. Rev. B **53**, p. 12694 (1996); A.P. Horsfield, A.M. Bratkovsky, D.G. Pettifor, M. Aoki, ibid., **53**, p. 1656 (1996).
8. A. Girschik, A.M. Bratkovsky, D.G. Pettifor, and V. Vitek, Phil. Mag. A, (in print), 1997.
9. Yang-Tse Cheng and Wen-Jin Meng, Phys. Rev. Lett. **76**, 3999 (1996).
10. N.W. Ashcroft and N.D. Mermin, Solid State Physics, Saunders College Publishing, London 1976, p. 759.

STRUCTURAL DISORDER AND LOCALIZED GAP STATES IN SILICON GRAIN BOUNDARIES FROM A TIGHT-BINDING MODEL

F. CLERI[1], P. KEBLINSKI[2], L. COLOMBO[3], S.R. PHILLPOT[2], D. WOLF[2]
[1]Divisione Materiali Avanzati, ENEA, Centro Ricerche Casaccia, C.P. 2400, 00100 Roma (Italy)
[2]Materials Science Division, Argonne National Laboratory, Argonne, IL 60439 (USA)
[3]Dipartimento di Scienza dei Materiali, Università di Milano, and Istituto Nazionale per la Fisica della Materia, via Emanueli 15, 20126 Milano (Italy)

ABSTRACT

Tight-binding molecular dynamics simulations of typical high-energy grain boundaries in silicon show that the atomic structure of the interface in thermodynamic equilibrium is similar to that of bulk amorphous silicon and contains coordination defects. The corresponding electronic structure is also amorphous-like, displaying extra states in the forbidden gap mainly localized around the coordination defects, where large changes in the bond-hybridization character are observed. It is proposed that such coordination defects in disordered high-energy grain boundaries are responsible for the experimentally observed gap states in polycrystalline Si.

INTRODUCTION

It is well known that the electrical properties of polycrystalline Si are dominated by localized electronic states in the band gap closely connected with the presence of grain boundaries (GBs). Such gap states can act as traps for charge carriers, thereby altering the electrical nature of the material. For example, in the case of p-n junctions for solar cells [1] the GBs provide effective recombination centers for the light-generated charge carriers, thereby reducing device performance [1,2]. Optical and electronic measurements on fine-grained Si films indicate the presence of exponential band tails penetrating into the band gap and narrow continua of states located deeper in the band gap [3-5]. The band tails are thought to be connected with structural disorder [6], through either the smearing of bulk states or the appearance of "shallow" gap states (*i.e.*, new states at the band edges). The deep gap states, in turn, are usually attributed to either dangling bonds (*i.e.*, 3-fold coordinated Si atoms) or to segregated impurities. Unfortunately, the fact that polycrystalline Si always contains sizable amounts of hydrogen and other impurities is a significant difficulty for the experimental investigation of the effect of GBs on the electronic properties. The structural and electronic properties of ideally pure GBs are, in turn, a subject particularly suitable for theoretical investigation by computer simulations.

Recent empirical-potential molecular dynamics (MD) simulations of several large-unit-cell Si GBs, either grown from the melt or annealed at high temperature, and subsequently cooled to T=0 K [7], revealed a common, highly disordered GB structure of uniform thickness (typically ~5 Å) and energy. Such disordered structure represents the thermodynamic ground state, as far as the high-temperature-annealed structure is found to have a lower GB excess energy with respect to the zero-temperature relaxed input structure of these same GBs. Moreover, the universal, "confined amorphous" structure of these GBs was found to be very similar to that of bulk amorphous Si (*a*-Si); in particular, a small fraction of coordination defects was found to exist *in thermal equilibrium* in this structure.

Previous electronic-structure simulations of Si GBs based simply on zero-temperature static relaxation have mainly focused on symmetric-tilt boundaries (STGB) [8-10], known to have a

relatively low energy compared to large-unit-cell, high-angle twist GBs [11]. The main result of these studies is that STGBs are relatively ordered and hence do not contribute deep gap states to the DOS. Similar simulations of twist GBs in Si [12] and Ge [13] revealed, by contrast, more disordered structures, containing coordination defects and exhibiting a broad continuum of gap states. (It is worth noting that, due to computational limitations, the only system studied was the the so-called $\Sigma 5$ (001) twist GB, see below.) It therefore appears that, although the highly disordered, high-energy GBs represent only a minor component of a well-annealed coarse-grained microstructure, their pronounced electrical activity may dominate the electrical behavior of rather pure polycrystalline Si. Such a dominance should become even more pronounced in fine-grained (10-50 nm) microstructures, in which a much larger fraction of high-energy GBs should be present [14]. Electrical measurements have, indeed, demonstrated a significant increase in the density of gap states with decreasing grain size [4]. These observations suggest that, withour explicitly simulating a random microstructure, insights into the electrical behavior of polycrystalline Si could be obtained from a deeper understanding of the properties of *high-energy* GBs.

TWIST GRAIN BOUNDARIES IN SILICON AND SIMULATION METHODS

As pointed out in earlier studies [7], from a purely geometric point of view the twist GBs on the (001) plane of the Si diamond lattice are representative of virtually all high-energy GBs, since two out of four bonds of each GB atom involve partners across the interface, while the remaining two bonds are within the same grain. This results in relatively high GB energies, compared to the GBs on the two densest planes (*i.e.*, (011) and (111)), in which only one bond per atom is directed across the interface. Based on these findings [7], we choose the (001) $\phi = 43.6°$ (so-called $\Sigma 29$) twist boundary as a typical high-energy GB for our high-temperature annealing TB-MD simulations; this GB has a planar unit-cell $\Sigma = 29$ times larger than the primitive planar unit cell of perfect-crystal (001) planes. For the sake of comparison, similar simulations will be performed also on the (001) $\phi = 36.9°$ ($\Sigma 5$) twist GB; this system was already studied in Refs. [12,13] however without taking into account the fundamental role of high-temperature annealing. As it will be shown, upon high-temperature annealing the $\Sigma 5$ GB displays all the basic features of the $\Sigma 29$ GB; moreover, its smaller planar unit cell will enable us to expose more clearly the correlation between atomic and electronic structure.

In all our simulations we consider a three-dimensionally periodic supercell with 24 (001) planes, containing two identical GBs separated by 12 (001) planes. The z-dimension of the simulation cell is then equal to $L_z = 6a_0$, while the x- and y-dimensions are equal to $L = (29/2)^{1/2} a_0$ and $L = (5/2)^{1/2} a_0$ for the $\Sigma 29$ and $\Sigma 5$ supercells, containing 696 and 120 atoms, respectively; $a_0 = 5.43$ Å is the Si lattice parameter. To describe the covalent bonding in Si we use an orthogonal TB Hamiltonian based on the minimal sp^3 basis set [15]; this TB parametrization is known to give a good description of Si point-defect properties [16]. Although less accurate than a fully *ab-initio* model, the TB formalism is nevertheless capable of describing in a simple and intuitive way the physics of covalent bonding and rehybridization in the presence of disorder, thus capturing the basic correlation between atomic structure and electronic properties. Moreover, it should be noted that MD simulations of such large systems as the $\Sigma 29$ GB are still out of reach for *ab-initio* techniques.

The input supercell configurations for the two GBs were obtained by high-temperature equilibration followed by cooling to, and relaxation at, zero temperature [7] using Tersoff's empirical potential [17]. These configurations were then subjected to TB-MD annealing at T=700 K, followed by cooling to, and relaxation at, T=0 K with TB quantum-mechanical forces.
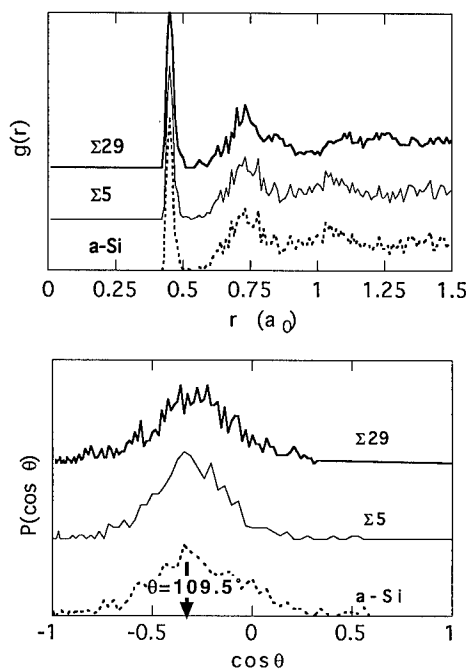
Figure 1. Pair distribution function (upper panel) and bond-angle distribution function (lower panel) averaged over the atoms in the planes adjacent the Σ29 and Σ5 grain boundaries. Dashed lines are the corresponding distribution functions for a-Si. The arrow in the lower panel indicates the crystalline-Si tetrahedral angle of θ=109.5°.

The relaxation was terminated when the largest force was smaller than 0.03 eV/Å. Stress relaxation in the z-direction perpendicular to the GB plane was allowed in both the Tersoff and TB cycles while keeping the xy-plane dimensions fixed so as to mimic embedding of the GB between bulk material.

ATOMIC STRUCTURE RESULTS

The results of TB-MD simulations confirm the previous empirical-potential results [7], thus supporting the notion that high-energy GBs in Si, indeed, exhibit a universal, confined-amorphous structure under thermodynamic-equilibrium conditions. The high-temperature-annealed Σ29 and Σ5 GBs display a disordered structure extending over a width of about 5 Å. Figure 1 shows the pair distribution function, g(r), and the bond-angle distribution function, P(cosθ), for the atoms comprised within the disordered region ($\pm a_0/2$ from the GB plane). Both sets of curves are very similar to the same distributions for a-Si (dotted line in Fig. 1), also

obtained with the procedure described above. These GBs have excess energies of E_{GB}=1864 mJ/m^2 and E_{GB}=1497 mJ/m^2, respectively. Both values are about 14% lower than the corresponding zero-temperature relaxed GB energies, in agreement with the value of about 10% energy lowering obtained by empirical-potential simulations on several high-energy GBs [7]. The average coordination in the GB regions is 4.05, somewhat lower than the empirical-potential value of 4.21 [7], but in good agreement with *ab-initio* [18] and TB [19] results for *a*-Si.

In the case of the Σ5 GB, we checked the stability of the high-temperature configuration against several others obtained from the conventional zero-temperature relaxation with TB forces. Each zero-temperature configuration is characterized by a different rigid-body translation vector in the GB plane, $t=(n_x,n_y)L/20$, with values of n_i=0, 1 or 2 to cover the irreducible part of the so-called "displacement-shift-complete" lattice [12]. The values of E_{GB} as a function of t cluster in two groups centered at about 1680 and 1920 erg/cm^2, with a spread of about 30 erg/cm^2 in each group. The minimum-energy configuration is found for t=(1,1): it has an excess energy E_{GB}=1664 erg/cm^2, *i.e.*, higher than that of the amorphous GB configuration by about 12%. Although considerably less disordered when compared to the amorphous GBs, all the zero-temperature-relaxed GB configurations display larger values of volume expansion and contain larger fractions of three-fold coordinated atoms. Some of these relaxed configurations were also subjected to the high-temperature cycle with the Tersoff potential and subsequent relaxation with the TB model: amorphous GB configurations were found which are statistically equivalent to the one described above, as far as the values of E_{GB}, average coordination, g(r) and P(cosθ) are concerned, although the details of the atomic structure could be locally different.

ELECTRONIC STRUCTURE RESULTS

We now turn to the study of the electronic properties of these confined-amorphous GBs. The electronic structure is calculated by sampling the **k**-space of the relaxed GB supercell, using the **k**=0 Γ-point for the Σ29 GB, and with a uniform mesh corresponding to about 2,000 points in the irreducible Brillouin zone of the diamond lattice for the Σ5 GB. The site- or orbital-projected local DOS is obtained from the eigenvector spectrum $c_{i\alpha}^n(\mathbf{k})$ by separately summing $|c_{i\alpha}^n(\mathbf{k})|^2$ over either the atomic index i or the orbital index α=s, p_x, p_y, p_z. Energies are measured with respect to the top of the valence band in the perfect crystal. The bottom of the conduction band (also defining the band gap) for this TB representation lies at 0.78 eV.

Figure 2 shows the local DOS averaged over the atoms in the disordered GB structure (top panel), and in the perfect-crystal-like planes (bottom panel) at z ~ ±a_0 from the GB, for both the Σ29 and the Σ5 GBs. By comparison with the *a*-Si DOS (also shown in the top panel of Fig. 2), obtained with the same **k**-space sampling density as above, it can be seen that the GB region is, indeed, amorphous-like also from an electronic-structure point of view. The local DOS shows the broadening of the valence s- and p-subbands (centered at about -10 eV and -4 eV, respectively, in the perfect crystal), characteristic of *a*-Si [20]; the conduction band is almost flat, in agreement with x-ray-absorption measurements on *a*-Si [21]. Only minor features of the perfect-crystal DOS persist in the GB local DOS, such as the small s-p mixing peak at about -6 eV. Several peaks in the band-gap region are visible; their origin will be discussed in more detail below. By contrast, the local DOS of atoms in the planes at z ~ ±a_0 is practically identical to that of crystalline Si, also shown in the bottom panel of Fig. 2, with only very minor signatures of the gap states; the latter are mostly due to finite-size effects.

516
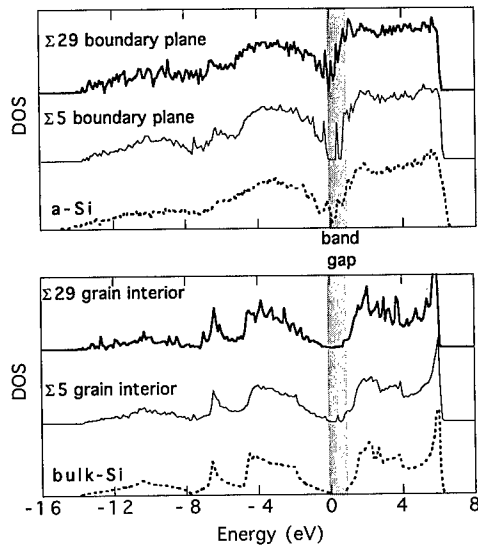
Figure 2. Upper panel: Electron DOS averaged over the atoms in the Σ29 and Σ5 grain boundaries; the dashed line corresponds to the a-Si DOS. Lower panel: Electron DOS averaged over the atoms in the planes at a distance $\pm a_0$ from the grain boundaries; the dashed line corresponds to the crystalline-Si DOS. The shaded region indicates the crystalline-Si band gap (0.-0.78 eV).

LOCALIZED GAP STATES AND COORDINATION DEFECTS

The correlation between the atomic structure of the GB and the gap states appearing in the GB local DOS can be elucidated by computing at each atom a localization index for the eigenvalues corresponding to the gap-state energies. This correlation can bee seen rather nicely already in the Σ5 GB. The localization index L_i^n for the n-th eigenvalue at atom i is defined as

$$L_i^n = \Sigma_\alpha |c_{i\alpha}^n|^2 \qquad , \qquad (1)$$

and gives the probability for atom i to contribute to that eigenvalue. L_i^n is thus equal to 1/N for a fully-distributed state in a supercell with N atoms, and equal to 1 if the state is entirely localized on atom i; we will consider a state to be at least partially localized whenever $L_i^n \geq 0.15$ for some i. In Figure 3 the energy-level arrangement of the gap states for all the atoms satisfying this localization criterion is displayed: not surprisingly, all the atoms labeled from 1 to 16 in Fig. 3 are located in the GB region. By summing L_i^n over the atoms in the GB for the eigenvalues
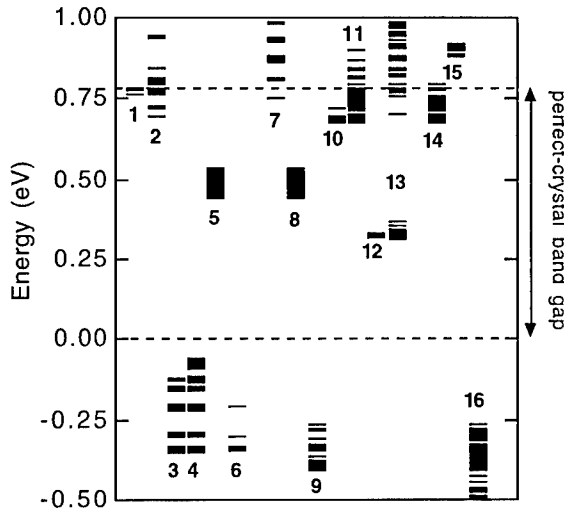
Figure 3. Energy-level distribution of the eigenvalues for which the localization index $L_i^n \geq 0.15$ in the $\Sigma 5$ GB. Each column of levels, numbered from 1 to 16, corresponds to a different atom i.

corresponding to the sharp peaks in the GB local DOS, values of $L^n \sim 0.5\text{-}0.7$ are obtained, indicating a distinctive localization of the gap states at the GB region.

A detailed analysis of the correlation between the localization index and the local atomic structure reveals that localized gap states in the electronic structure arise from coordination defects in the disordered atomic structure of the GB. Figure 4 shows the equilibrium configuration of this GB, in which all the atoms satisfying the localization criterion $L_i^n \geq 0.15$ are highlighted. The GB region contains one 3-fold- (atom 13) and some 5-fold-coordinated atoms (2, 5, 7, 8 and 12); moreover, some atoms participate in one or more odd-membered rings. All of the 3- and 5-fold, and some of the 4-fold coordinated atoms in the GB structure are found to display localized states; the local DOSes for three such atoms (one for each coordination state) are also shown in Fig. 4. From this analysis it appears that deep gap states are mostly localized at coordination defects, with very high values of the localization index (up to $L_i^n \sim 0.57$ for $E^n = 0.31$ eV). Shallow band-edge states, in turn, can be localized either at such defects or at their nearest-neighbors, with somewhat more distributed values of L_i^n. The only partial exception is represented by atoms 11 and 14, which give rise to shallow states around $E^n = 0.70\text{-}0.75$ eV: these atoms are the only ones not being neighbors to any coordination defect but are, however, in common to two seven-membered rings, i.e., another kind of topological disorder also typical of a-Si.

The nature of the gap states becomes apparent when analyzing the local bonding around coordination defects in terms of its s- and p-orbital components. The distribution of the band-
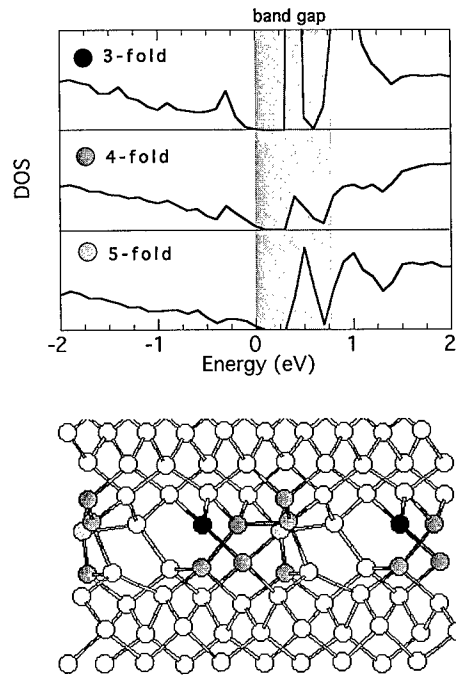
Figure 4. Upper panel: electron DOS for representative 3-fold, 4-fold and 5-fold coordinated atoms in the grain-boundary region; the shaded region indicates the crystalline-Si band gap. Lower panel: periodically-repeated equilibrium structure of the Σ5 grain boundary, with the atoms displaying localized gap or edge states highlighted (black=3-fold, dark grey=4-fold, light grey=5-fold coordinated atoms).

structure energy (*i.e.*, the sum of the single-particle energies for the occupied states), U_{BS}, has large variations around the coordination defects. Such variations, ΔU_{BS}, can be traced back to a change in the degree of s-p mixing, by calculating for each atom i the s-p mixing parameter [10],

$$M_i = N_{ip}/N_{is} \qquad (2)$$

with

$$N_{ip} = \Sigma_{n,\alpha} |c_{i\alpha}^n|^2, \quad \alpha = p_x, p_y, p_z$$

$$N_{is} = \Sigma_n |c_{n,\alpha=s}|^2 \qquad (3)$$

Perfect sp^3 hybridization would result in $M_i=3$, while $M_i=1$ is the value for a free Si atom with its valence electrons evenly distributed in unhybridized 3s and 3p orbitals. However, due to the competition between the hybridization energy and the promotion energy, we find $M_{dc}=1.684$ for atoms in bulk Si meaning that, even in the perfect crystal, bonds are never formed from complete sp^3 hybrids.

By plotting ΔU_{BS} vs. $\Delta M=(M_{dc}-M_i)/(M_{dc}-1)$ in Figure 5, we find a good correlation between the variation in the covalent-bond energy and the change in the hybridization character. The largest s-p mixing variation is observed for the three-fold defect (atom 13), whose $\Delta M=0.47$ reflects a drastic change towards a s^2p^2-like hybridization [10], the dangling electron tending to form states with a large s-orbital component; all those five-fold defects which display deep gap states (atoms 5, 8 and 12) also have sizable variations of the hybridization character, $\Delta M=0.22$-0.26; the four-fold coordinated atoms 6 and 15, which appear to be neighbors to two coordination defects at once, display variations of ΔM in this same range and very large values of ΔU_{BS}. Clearly, such large changes in the s-p hybridization character around coordination defects are responsible for the appearance of gap states at energies which are extraneous to tetrahedrally-bonded Si.
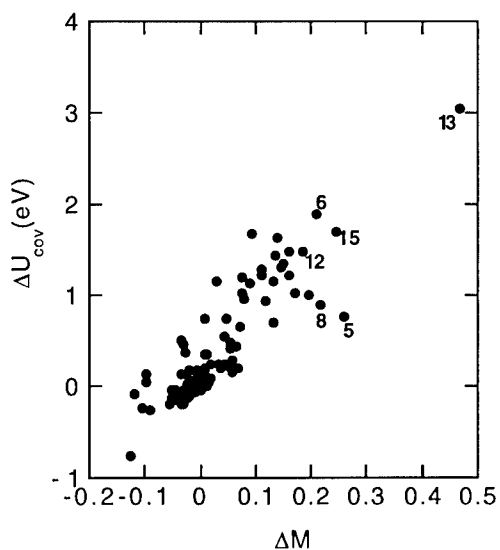
Figure 5. Correlation plot of the band-structure energy variation, ΔU_{BS}, vs. the s-p mixing character variation, ΔM_i.

CONCLUSIONS

The main conclusion of our study is that gap states in *pure* polycrystalline Si can originate from coordination defects present in the amorphous-like structure of the high-energy GBs. It is noteworthy that a similar correlation between coordination defects and localized gap states had previously been found in a detailed theoretical study of *a*-Si [22]. Despite this, the connection between the localization behavior of coordination defects in *a*-Si and localized gap states in polycrystalline Si has gone practically unnoticed. This is even more surprising, when considering that amorphous-like behavior of GB regions had been repeatedly suggested, *e.g.*, in the analysis of photoconductivity of undoped Si [23], of optical-absorption measurements on columnar polycrystals [3], and of the DOS obtained from field-effect conductance in fine-grained samples [4]. Notably, Paxton and Sutton [11] discounted coordination defects in favor of impurity segregation as a source of gap states, "unless a mechanism of producing and sustaining a non-equilibrium distribution of these defects can be devised". We have shown that such a mechanism can indeed be provided by the amorphous GB film, however with the important difference that the defect distribution is here *in thermodynamic equilibrium*. Although impurities may also play an important role, the above results demonstrate that gap states in polycrystalline Si, arising from coordination defects existing in thermodynamic equilibrium, are an intrinsic feature of the system arising from the universal, disordered structure of the high-energy GBs.

Upon hydrogen contamination the 3-fold defects (dangling bonds) are preferentially passivated [3-5], while the 5-fold defects are much less affected. (Some experiments [5] indicate that the deep states surviving hydrogenation are shifted to higher energies, which is qualitatively visible in our Fig. 3 where the 5-fold-defect gap states lie about 0.2 eV above the 3-fold ones.) Because coordination defects carry some fraction of excess charge, they can realize *intrinsic* accumulation/depletion layers along the amorphous GBs [1]. This possibility warrants further investigation by a more accurate *ab-initio* technique, capable of elucidating also the charge density distribution around coordination defects.

ACKNOWLEDGEMENTS

REFERENCES

1. L.M. Fraas, J. Appl. Phys. 49, 871 (1978).
2. K.V. Ravi, *Imperfections and Impurities in Semiconductor Silicon*, (John Wiley, New York, 1981), p.298.
3. W.B. Jackson, N.M. Johnson and D.K. Biegelsen, Appl. Phys. Lett. 43, 195 (1983).
4. G. Fortunato and P. Migliorato, Appl. Phys. Lett. 49, 1025 (1986).
5. D. Jousse, S.L. Delage and S.S. Iyer, Phil. Mag. B63, 443 (1991).
6. C.M. Soukoulis, M.H. Cohen and E.N. Economou, Phys. Rev. Lett. 53, 616 (1984).
7. P. Keblinski, S. R. Phillpot, D. Wolf and H. Gleiter, J. Am. Cer. Soc. 80, 717 (1997).
8. D.P. DiVincenzo, O. L. Alerhand, M. Schluter and J. W. Wilkins, Phys. Rev. Lett. 56, 1925 (1986).

9. M. Kohyama and R. Yamamoto, J. Phys. C, Solid St. Phys. 21, 3205 (1988).
10. A.T. Paxton and A.P. Sutton, Acta Metall. 37, 1693 (1989).
11. D. Wolf, J. Physique C4, 197 (1984).
12. M. Kohyama and R. Yamamoto, Phys. Rev. B49, 17102 (1994).
13. E. Tarnow et al., Phys. Rev. B42, 3644 (1990).
14. P. Keblinski et al., Acta Mater. 45, 987 (1997).
15. I. Kwon, R. Biswas, C. Z. Wang, K. M. Ho and C. M. Soukoulis, Phys. Rev. B49, 7242 (1994).
16. M. Tang et al., Phys. Rev. B55, 14279 (1997).
17. J. Tersoff, Phys. Rev. Lett. 61, 2879 (1988).
18. I. Stich, R. Car and M. Parrinello, Phys. Rev. B44, 11092 (1991).
19. L. Colombo and M. Maric, Europhys. Lett. 29, 623 (1995).
20. L. Ley, S. Kowalczyc, R. Pollack and D. A. Shirley, Phys. Rev. Lett. 29, 1088 (1972).
21. J. Reichardt, L. Ley and R.L. Johnson, J. Non-Cryst. Solids 35/36, 256 (1985).
22. P.A. Fedders, D.A. Drabold and S. Klemm, Phys. Rev. B45, 4048 (1992).
23. S. Hasegawa, S. Takenaka and Y. Kurata, J. Appl. Phys. 53, 5022 (1982).

ELECTRONIC STRUCTURE OF AMORPHOUS SILICON

G. ALLAN, C. DELERUE, M. LANNOO
Institut d'Electronique et de Microélectronique du Nord, Département Institut Supérieur d'Electronique du Nord, BP 69, 59652 Villeneuve d'Ascq Cedex, France, gal@isen.fr

ABSTRACT

The electronic structure of a continuous network model of tetrahedrally bonded amorphous silicon (a-Si) and of a model hydrogenated amorphous silicon (a-Si:H) that we have built from the a-Si model are calculated in the tight binding approximation. The band edges near the gap are characterized by exponential tails of localized states induced mainly by the variations in bond angles. The spatial localization of the states is compared between a-Si and a-Si:H. Valence band offset between the amorphous and the crystalline phases is calculated.

INTRODUCTION

Recently, large (4096-atom supercell) computer-generated continuous network models of tetrahedrally bonded amorphous carbon and silicon have been constructed such as the Wooten-Winer-Weaire (WWW) model [1-2]. This one reproduces with good accuracy the experimental radial distribution function of amorphous silicon (a-Si) and for a smaller unit cell (216-atom model) calculations of the first-order Raman spectrum are in very good agreement with the experimental one [3], which, as a test to measure the credibility of structural models [4], gives confidence that the WWW a-Si model is a realistic one. We have calculated its electronic structure in the tight-binding approximation. This exhibits a large number of deep and strongly localized states which correspond to the well known band tails in the amorphous semiconductor band gap. Such a large band tailing is attributed to a small number of bond angles which deviate greatly from the tetrahedral 109° value.

In the next section, we describe the a-Si:H supercell created by hydrogenation of the a-Si atomic WWW model. Then we calculate the electron density of states which shows a small increase of the band width between amorphous and crystalline silicon (c-Si). The slope of the band gap tails are also determined and decrease from a-Si to a-Si:H. Finally the valence band offsets between a-Si, a-Si:H and c-Si is calculated.

THE MODEL

A complete total energy calculation which would give the positions where hydrogen atoms are in a-Si:H is not yet feasible for large supercells. Starting from the WWW model, we have assumed that the minimum energy configuration is obtained when the hydrogen atoms are inserted in the regions of large bond angle distorsions (the bond length variations are quite small in a-Si). As there is a direct relation between the bond angle distorsions and the a-Si localized states, we have first determined the Si atoms which give rise to these gap states.

For this, we "cut" small clusters (with about 50 Si atoms) centered on every atom of the supercell and saturate the dangling bonds of each cluster with hydrogen atoms. Then the highest occupied molecular orbital (HOMO) and the lowest unoccupied one (LUMO) are calculated in the tight-binding approximation as described below. When we look at the clusters with the smallest gaps, the corresponding HOMO and LUMO states are found mainly localized on the same atom with strong bond distortions. To simplify the hydrogenation process we have removed the central atoms one by one and the four dangling bonds created by the atom removal are saturated with

hydrogen atoms as in Ref. 5. The final concentration of hydrogen atoms is equal to 8% (83 silicon atoms were removed). Once this is done the atomic positions and the supercell lattice parameter are again relaxed by use of a Keating potential [6] as in the original WWW model. This procedure is rather empirical but we expect that it gives a good relaxation of the strain in the supercell.

The electronic structure of the a-Si and a-Si:H supercells has been calculated in the tight-binding sp^3s^* Vogl's approximation [7]. Even if it is a first-nearest neighbor model, the bulk c-Si band gap which is fitted is in good agreement with the experimental values. The model cannot be used to accurately determine states high in the conduction band as the resulting c-Si conduction bands are rather flat [8]. This is not too important in this calculation as the conduction band edge will always remain close to bulk c-Si one. We take a usual d^{-2} variation of the tight-binding parameters with the interatomic distance d [9]. Harrison's parameters [9] are used for the first-nearest neighbor Si-H interaction parameters. Only the Γ point at the center of the supercell cubic Brillouin zone has been used since, due to the large size of the supercell, we expect a negligible dispersion of the energies versus wave vector.

THE DENSITIES OF STATES

The local densities of states on each atom of the supercell are approximated by a continued fraction whose 40 first coefficients are calculated by the recursion method [10]. A semi-elliptic termination is used and their values are determined by a linear prediction method [11]. Such a procedure uses 80 exact moments of the density of states which is sufficient to get converged results [12].

Figure 1 shows the average density which is close to those obtained in preceding calculations [13]. The amorphous silicon valence band width is slightly larger than the c-Si one. This seems to be due to the variation of the interatomic distances in the amorphous phase as the shift of the bottom of the band (~0.3 eV) disappears when the variation of the interatomic tight-binding parameters with the distance is suppressed. The main difference between a-Si and a-Si:H occurs near the
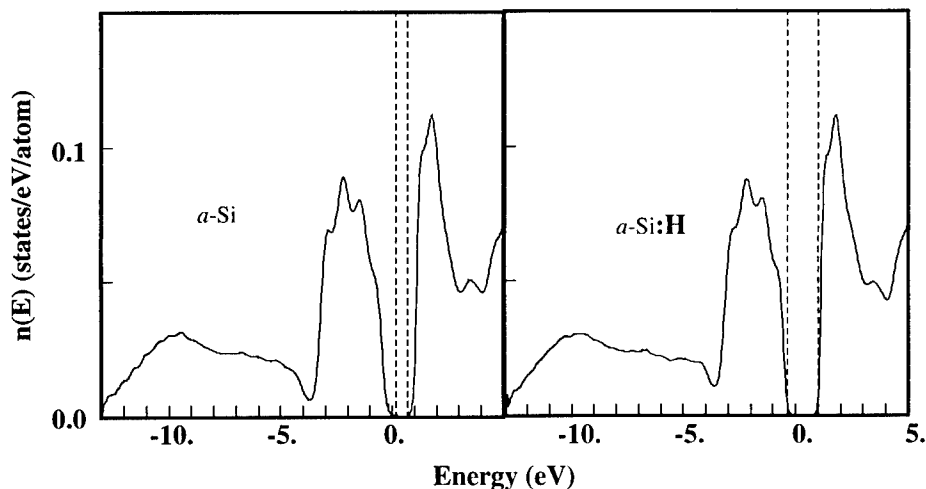
Average density of states of a-Si (a) and a-Si:H (b) as calculated by the recursion method. The dotted lines indicate the limits of the gap.

band gap. We see that the «hydrogenation» has suppressed the deep states in the gap.

To get a more precise information about these states, one has to diagonalize the hamiltonian H. For this, we have used a minimization of the Rayleigh quotient $\langle \Psi | (H - \sigma I)^2 | \Psi \rangle / \langle \Psi | \Psi \rangle$ with respect to a trial function Ψ [14]. σ is an energy shift close to the band gap .

As the supercell is finite, the limits of the band gap are well defined and are equal to (-0.36 eV, 1.00 eV) for a-Si:H and (0.19 eV, 0.72 eV) for a-Si. We shall see below that the a-Si band edges correspond to states strongly localized on a few atoms whereas in the a-Si:H case, they correspond to moderately localized states. One must notice that these values do not depend very much on the variation of the interatomic distances. For example, the top of the a-Si:H valence band is simply shifted by 0.04 eV when the variation of the tight-binding parameters with interatomic distance is neglected and when all are taken equal to the c-Si ones. This shows that the variation of the band gap limits is mainly due in the amorphous phases to bond angle distortions.

The band lineups between crystalline and amorphous silicon can be obtained from these gap values if one knows the electrostatic dipole layer at an heterojunction between these materials. In the molecular model and in the zero charge transfer approximation which has been often used for semiconductor heterojunctions [15], the electrostatic dipole layer is just equal to the opposite of the energy difference between the hybrid orbitals on silicon atoms close to the heterojunction. The average bond angle for the a-Si:H model which is equal to 109°27 is very close to the c-Si one which is determined by the 109°47 tetrahedral bond angle so that the mean hybrid orbital energy in a-Si will be close to the c-Si. Using c_s and c_p the average s and p characters for the valence band that one can calculate using the recursion method ($c_s + c_p = 4$), one can define an average hybrid orbital $(c_s E_s + c_p E_p)/4$ for c-Si, a-Si and a-Si:H where E_s and E_p are the s and p Si orbital energies, respectively. The energy difference between the amorphous and crystalline silicon is less than 0.02 eV. So the electrostatic dipole layer at the heterojunction between these materials is negligible. The valence band offset is thus directly obtained from the limits of the band gap discussed before. It is then equal to 0.36 eV between a-Si:H and c-Si (with a valence band higher in c-Si) and equal to -0.19 eV between a-Si and c-Si (with a valence band higher in a-Si) in agreement with recent *ab initio* pseudopotential calculation for a smaller supercell [16]. For a-Si:H, this is also in agreement with the experimental results which vary from 0 to 0.71 eV [16] depending on preparation conditions of the amorphous materials.

The density of states due to localized states in amorphous semiconductors is often assumed to decay exponentially away from the conduction and valence band edges [17-19]. This would also be true for the integrated density. Figure 2 shows the fit of the staircase integrated density by exponential curves $\exp(-|E|/E_0)$ where E_0 is the band tail slope. The slopes E_0 are equal to 142 meV and 81 meV respectively for the a-Si valence and conduction bands. These values are reduced to 37 meV and 19 meV for a-Si:H. The values for a-Si:H are slightly smaller than the experimental ones which are in the range of 43-103 meV (for the valence band tail) and 27-37 meV (for the conduction one) [17-21]. Let us recall that due to the finite size of the supercell, we do not have very large wavelength lattice distortions. Their effect will be to spread the band tails we have obtained and then to slightly increase the slope values we have calculated.

a-Si AND a-Si:H STATES LOCALIZATION

We shall characterize this localization by:

$$N_{at}(E_j) = \left(\sum_i \left(\sum_{l=1}^{5} a_{il}^2 \right)^2 \right)^{-1} \tag{1}$$
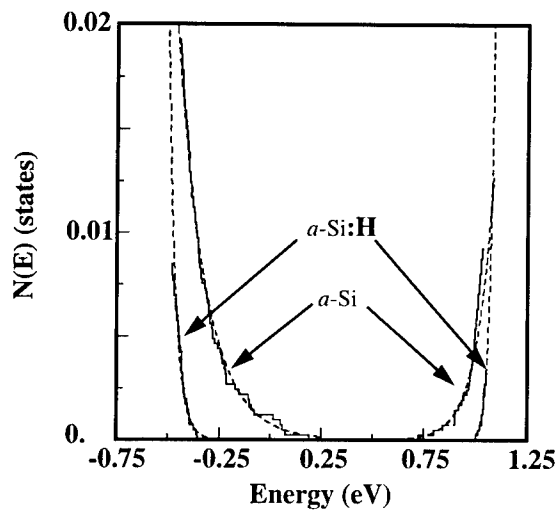
Fig. 2 Integrated density of states near the band gap (full line). The exponential fit (dotted line) is almost indistinguishable from the staircase integrated density.
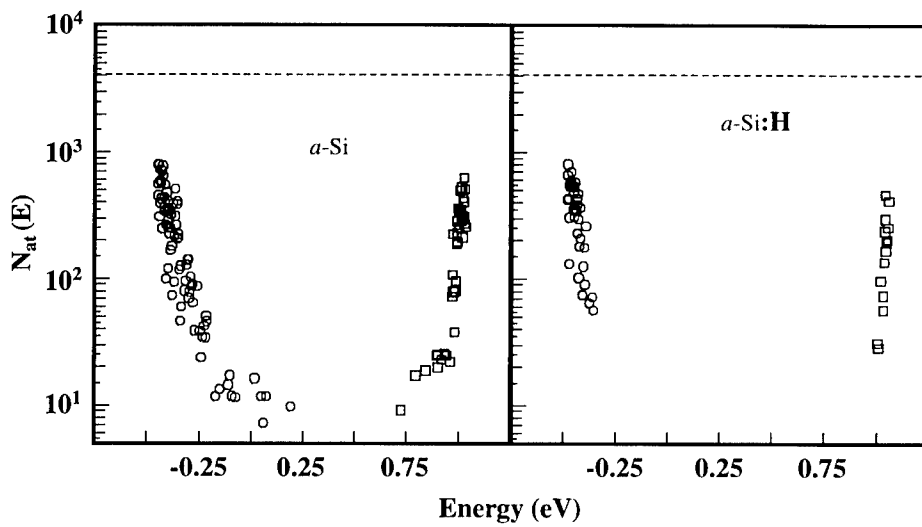
Figure 3 Spatial localization of the band tail states. The dotted line indicates the number of Si atoms in the supercell.

where the sum is extended to all the supercell atoms $\vec{R}_i$ and to the 5 sp^3s^* atomic functions $\phi_l(\vec{r} - \vec{R}_i)$, the tight-binding eigenfunction for the energy E_j being equal to:

$$\psi_j = \sum_{i,l} a_{il}\phi_l(\vec{r} - \vec{R}_i) \tag{2}$$

When an eigenstate is completely uniform over n atoms, we have $\sum_{l=1}^{5} a_{il}^2 = 1/n$ and $N_{at}(E_j)$ is just equal to n. So $N_{at}(E_j)$ gives an order of magnitude of the number of atoms on which the state is localized. Figures 3 shows this localization for the states close to the gap. For a-Si, states close to the band edges are found to be localized on about 10 atoms. Such states are suppressed by the hydrogenation. Nevertheless for a-Si:H, the gap edges are still localized on a few tens of atoms. This number rapidly increases as one goes deeper for example in the valence band but 0.15 eV below the valence band edge, the states still extend over less than a thousand atoms.

CONCLUSION

The electronic structure of large amorphous silicon cells has been calculated in the tight-binding approximation. The width of the band is slightly increased due to the small variations of the interatomic distances. Localized gap states which are due to bond angle distortions have been used to create an atomic a-Si:H model. The a-Si and a-Si:H valence band offset and the slopes of the exponential tails of the density of states close to the gap are in good agreement with experimental values. This gives good confidence in the structural models.

ACKNOWLEDGEMENTS

REFERENCES

1 F. Wooten, K. Winer, and D. Weaire, Phys. Rev. Lett. **54**, 1392 (1985); F. Wooten, and D. Weaire, Solid State Phys. **40**, 1 (1987).
2 B. R. Djordjevic, M.F. Thorpe, and F. Wooten, Phys. Rev. B **52**, 5685 (1995).
3 M. Marinov, and N. Zotov, Phys. Rev. B **55**, 2938 (1996).
4 N. Maley, D. Beeman, and J. S. Lannin, Phys. Rev. B **38**, 10611 (1988).
5 P. A. Fedders, and D. A. Drabold, Phys. Rev. B **47**, 13277 (1993).
6 P. N. Keating, Phys. Rev. **145**, 637 (1966).
7 P. Vogl, H. P. Hjalmarson, and J. D. Dow, J. Phys. Chem. Sol. **44**, 365 (1983).
8 C. Delerue, M. Lannoo, and G. Allan, Phys.Rev.Lett. **76**, 3038 (1996).
9 W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, San Francisco, 1980).
10 R. Haydock, *Solid State Phys.* **35** (Academic Press, New York, 1980) p. 216.
11 G. Allan, J. Phys. C (Solid State Physics) **17**, 3945 (1984).
12 J. Dong, and D. A. Drabold, Phys. Rev. B **54**, 10284 (1996).
13 J.M. Holender,G.J. Morgan, J. Phys. C (Solid State Physics) **4**, 4473 (1992)

14 L. W. Wang and A. Zunger, J. Chem. Phys. **100**, 2394 (1994).

15 C. Priester, G. Allan, and M. Lannoo, J. Vac. Sci. Technol. B **6**, 1290 (1988).

16 C. G. Van de Walle, and L. H. Yang, J. Vac. Sci. Technol B **13**, 1635 (1995).

17 T. Tiedje, J. M. Cebulka, D. L. Morel, and B. Abele, Phys. Rev. Lett. **46**, 1425 (1981).

18 K. Winer, I. Hirabayashi, and L. Ley, Phys. Rev. Lett. **60**, 2697 (1988).

19 S. Aljishi, J. D. Cohen, S. Jin, and L. Ley, Phys. Rev. Lett. **64**, 2811 (1990).

20 C.-Y. Huang, S. Guha, and S.J. Hudgens, Phys. Rev. B **27**, 7460 (1983).

21 T. Muschik, R. Schwarts, and F. Karg, J. Lumin. **48-9**, 636 (1991).

CARBON SCHWARZITES: PROPERTIES
AND GROWTH SIMULATION FROM FULLERENE FRAGMENTS

G. BENEDEK[1,2], L. COLOMBO[1,2], S. SPADONI[1,2], S. GAITO[1,2], and P. MILANI[1,3]

[1] Istituto Nazionale per la Fisica della Materia (INFM), UdR Milano–Università
[2] Department of Materials Science, via Emanueli 15, I-20126 Milano, Italy
[3] Department of Physics, via Celoria 16, I-20133 Milano, Italy

Abstract

We investigated by tight-binding molecular dynamics the structure, the bulk modulus and the electronic properties of the smallest four-branched carbon schwarzites fcc-$(C_{28})_2$, fcc-$(C_{36})_2$ and fcc-$(C_{40})_2$ having the shape of a D minimal periodic surface. They are found to have a stability comparable to that of fullerene C_{60} and to exhibit alternative metallic and insulating characters, with an apparent relationship to their local geometry. We also studied the coalescence of fullerenic fragments and carbon clusters by following the evolution of the topological connectivity. Though different temperature variation protocols lead to irregular structures similar to random schwarzites, their connectivity is found to stabilize at values corresponding either to tubulene or three-branched schwarzites, indicating that long time evolution at constant connectivity is potentially able to yield regular shapes. Experiments on laser-induced transformations of fullerite occasionally yield branched tubular structures with a schwarzite shape.

Introduction

One of the most challenging development in modern technology is represented by devices based on atomic-scale porous materials. The interesting features that such materials exhibit range from large specific area (that permit to intercalate a great deal of lighter atoms such as alkali) to a large stability and stiffness (to prevent induced stress during the intercalation, unlike graphite that is damaged when intercalated.) In this framework carbon schwarzites [1] are predicted to be very interesting and promising materials for technological applications, e.g. as lithium absorbers for various functions in ionic devices. In particular, conducting schwarzites could be employed in cathodes (or/and anodes, depending on the combination of chemical potentials), while insulating schwarzites may work as low-temperature ionic conductors or molecular sieves.

Most of schwarzite intriguing features derive from their peculiar atomic structure. The topology of schwarzites can be constructed solving the Euler's theorem for surface polygonal tilings. The theorem has to be referred to a single element, since schwarzites are open and infinitely extended surfaces.

Among all the possible minimal periodic surfaces we constructed the topology of the D-type minimal periodic surfaces where each element coordinates four identical elements in the tetrahedral (staggered) configuration. The resulting crystal is fcc and has, like diamond, two elements per unit cell. In order to find the connectivity of this kind of schwarzites we can imagine to close a unit cell on itself by joining three pairs of opposite branches and transforming the surface in the topological equivalent of a three-hole torus of connectivity 7. Therefore, the Euler's theorem can be recast in the following form:

$$v_{el} - e_{el} + f_{el} = 3 - K = -2 \tag{1}$$

where v_{el}, e_{el} and f_{el} are the numbers of vertices (atoms), edges (bonds) and polygonal faces (rings) per element and K is the connectivity. We further focused on schwarzites with only heptagons and hexagons because they are the equivalent of fullerenes. Indeed by solving the Euler's theorem with these conditions we find that it is necessary to use 12 heptagons and an arbitrary numbers of hexagons compatibly with the symmetry of the crystal. Morever it is possible to use topology to characterize the evolution of a system of atoms by its order of connectivity[2]. From Euler's theorem we can compute the connectivity from the number (f_k) of k-membered rings of the systems:

$$K = -\sum_k f_k + e - \sum_n v_n + 3 \tag{2}$$

where $e = \sum_n (n/2)v_n$ is the total number of edges (bonds) in the simulation cell and v_n is the number of n-coordinated atoms.

Structural, elastic and electronic properties of fcc-$(C_{28})_2$, fcc-$(C_{36})_2$ and fcc-$(C_{40})_2$.

We studied the three smallest schwarzites of the series made with only heptagons and hexagons [3]. The smallest one has only heptagons formed by 28 atoms per unit cell (Fig.1) and shows a chirality; the second one contains four hexagons in each element lying on the faces of a tetrahedron and has 36 atoms per element; the largest schwarzite here studied has six hexagons lying on the edges of a tetrahedron and its element is composed by 40 atoms. We calculated the equilibrium structure, the ground-state properties and the electronic structure of these schwarzites by performing tight-binding molecular dynamics (TBMD) simulations with the parametrization of Xu et al.[4].
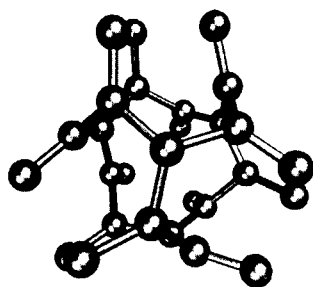
Figure 1: The unit cell of the schwarzite fcc-$(C_{28})_2$.

This scheme conjugates a low computational cost as needed for simulating the present systems to an accurate quantum-mechanical treatment as required to study the chemistry of the C-C bond (i.e. its allotropy, directionality and covalence)[5].

We obtained the equilibrium structure by a simultaneous minimization of both the total energy (performed by TBMD relaxations at T=0 K and with cubic cells containing respectively 288, 300 and 320 atoms), and of the density (using a Murnagham-like equation of state). Canonical simulations at room temperature have confirmed the stability of the structures.

In Table I we report the calculated cohesive energy per atom (relative to diamond), the density, the bulk modulus (B) and the stiffness per bond (b), defined as

$$b = \frac{2}{3}BV_{at} \qquad (3)$$

where V_{at} is the atomic volume.

Table 1: Calculated cohesive energy with respect to diamond (E_{coh}), density, bulk modulus (B), stiffness per bond (b) and electronic behavior of fcc-$(C_{28})_2$, fcc-$(C_{36})_2$ and fcc-$(C_{40})_2$. The corresponding values for fullerite and diamond (as obtained with the present TBMD model) are reported for comparison.

	E_{coh} (eV/atom)	Density (g/cm³)	B (Mbar)	b (MbarÅ³)	
fcc-$(C_{28})_2$	0.70	1.33	1.58	16.12	Metal
fcc-$(C_{36})_2$	0.65	1.05	1.26	16.20	Insulator
fcc-$(C_{40})_2$	0.44	1.60	1.92	16.25	Metal
Fullerite	0.37	1.71			Insulator
Diamond		3.52	4.42	16.71	Insulator

It is apparent from Table I that these new crystals have a cohesive energy similar to that of fullerene even if they are the smallest of the series and consequently the less energetically favored. Moreover they have a density compared to either fullerite or graphite–a remarkable property related to the fact that, unlike graphite and fullerite, schwarzites are fully covalent in three dimensions and have a stiffness per bond similar to that of diamond. Both the band structure and the electronic density of states (DOS) have been calculated with the same tight-binding parametrization [3]. These crystals exhibit alternative metallic and insulating behaviors depending on their geometry.

Characterization of the structural evolution during and after the coalescence of carbon fragments.

A natural question is whether it is possible to syntetize schwarzites. We started from the experimental work of Milani et al.[6] e Ferretti et al.[7] who have recently used laser-induced transformations of fullerite to process carbon nano-structures even if it is very hard to predict the final atomic structure of the sample. We tried to understand the relationship among the nature of precursors in the disordered phase, the temperature effects and the topology of the final condensed phase. With this aim we studied in the same computational framework as above the early stages of the aggregation process of carbon fragments by monitoring the connectivity of the system as a function of time.
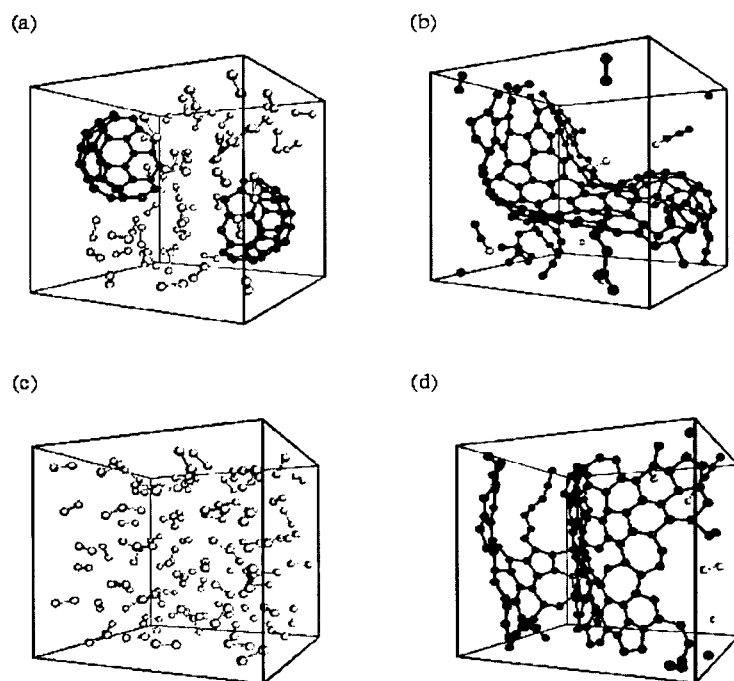
Figure 2: Snapshots of the initial (a) and final (b) configuration of a system of two C_{39} fragments embedded in a gas of C_2 dimers, after annealing through three different temperatures. Similar snapshots (c), (d) are shown also for a system of 90 C_2 dimers annealed at a constant temperature of 3500 K.

We performed three constant-temperature, constant-volume simulations differing for starting configurations and annealing procedure, but with the same number of atoms per cell (180 atoms) and the same density of the sample fixed at 0.876 g/cm^3 which is smaller

than solid fullerite, but higher than vaporized graphite as experimentally observed in laser ablation conditions. We choose a time step of 0.7 fs.

The first simulation was performed on a system of two C_{39} bowl-shaped cap fragments embedded into a gas of 51 C_2 dimers (Fig.2 (a)). The starting configuration was annealed consecutively at three different temperatures 1500, 2500, 3500K. The configuration reached at the end of this simulation is shown in Fig.2 (b) and appears to be very similar to a random schwarzite described by Townsend et al.[8]. A very interesting feature is the presence of 7-membered rings which introduce the local negative curvature characteristic of schwarzites. This schwarzite-like configuration is also confirmed by the value of the connectivity that stabilizes in $K=5$, as shown in Fig.3. In the second simulation, where the two fragments in a dimer gas have been annealed at one single constant temperature T=3500 K , K tends to a value much smaller then 5 and the average coordination is resulted to be 2.81 while greater than the 2.68 of the first simulation, indicating a more graphite-like behavior.

These two simulations have also demonstrated that the final configuration depends on the precursors, as well as on the thermal treatment.
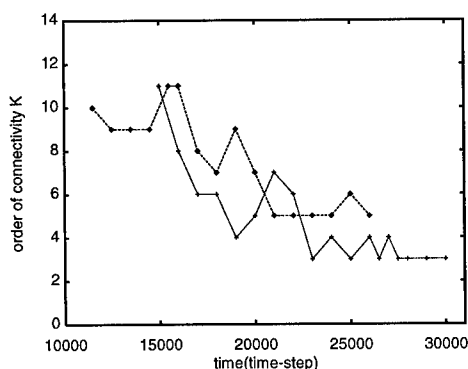
Figure 3: Time evolution of the order of connectivity K for the system of two C_{39} fragments (dashed line with $\diamond$ symbols) and the gas of 90 C_2 dimers (dotted line with + symbols).

In order to check the validity of our topological analisys we performed a third different simulation: a gas of ninety C_2 dimers (Fig.2 (c)) was annealed at T=3500 K for 35000 time steps. Now stabilization occurs at the value K=3 (Fig.4) that is the topological fingerprint of an open-ended tubulene easly visible in Fig.2 (d). The choice of this starting configuration is based on the lack of precursors for condensation that make the evolution of the system depending only on thermodynamics, i.e. on temperature and density conditions.

In conclusion these TBMD simulations have shown that under certain conditions, annealing experiments on fullerite may lead to branched tubulenes, or at least to topological structures with the appropriate topology for a continous transformation in a schwarzite. With different precursors and annealings, however, the system preferably evolves into graphite-like structures. From the observation that the order of connectivity, after an initial time of rapid transformations remains constant during the evolution of the condensed system, thus
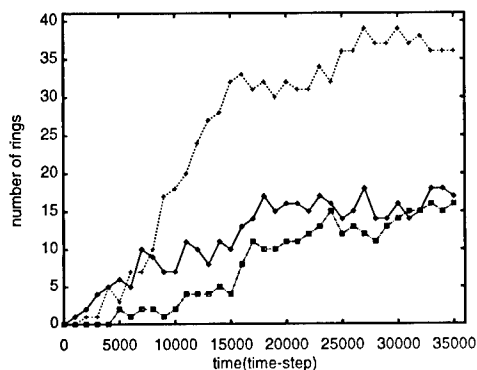
Figure 4: Time evolution of the ring statistics for the systems of 90 C_2 dimers shown in Fig.3 (c–d). Symbols: ◇ 5-membered rings; + 6-membered rings; □ 7-membered rings.

providing an important index to monitor and classify the possible structures towards which the system can evolve over longer time scales.

Acknowledgments

References

[1] A. L. Machay, Nature,**314**, 604 (1985); The term "schwarzite" honors the work of H. A. Schwarz, Gesammelte Mathematische Abhaudlungen, Vols. **1,2** (springer, Berlin) 1890, on the minimal periodic surfaces.

[2] S. Spadoni, L. Colombo, P. Milani and G. Benedek, Europhys. Lett. **39**, 269 (1997).

[3] S. Gaito, L. Colombo and G. Benedek, to be published.

[4] C. H. Xu, C. Z. Wang, C. T. Chan and K. M. Ho, J. Phys. Condens. Matter, **4**, 6047 (1992).

[5] L. Colombo, in "Annual Reviews of Computational Physics", Vol. IV, p.147, edited by D. Stauffer (World Scientific, Singapore, 1996).

[6] P. Milani and M. Manfredini, Appl. Phys. Lett., **68**, 1769 (1996).

[7] M. Ferretti, A. Parisini, M. Manfredini and P. Milani, Chem. Phys. Lett., **259**, 432 (1996).

[8] S. J. Townsend, T. J. Lenosky, D. A. Mueller, C. S. Nichols and V. Elser, Phys. Rev. Lett., **69**, 921 (1992).

AUTHOR E-MAIL INDEX

AUTHOR INDEX

SUBJECT INDEX